中国科学院教材建设专家委员会规划教材

供医学七年制、八年制、研究生使用

基因工程学

孙汶生
曹英林 主编
马春红



科学出版社
www.sciencep.com

中国科学院教材建设专家委员会规划教材

供医学七年制、八年制、研究生使用

基因工程学

主　编　孙汶生　曹英林　马春红
副主编　刘素侠　魏　然
编　者　（以姓氏笔画排序）
于学慧　马春红　王晓燕　王　群
石永玉　朱法良　刘素侠　孙汶生
宋　静　张利宁　赵　丽　郭　春
高立芬　高雪芹　曹英林　韩金祥
韩丽辉　魏　然

科学出版社
北　京

内 容 简 介

本书是山东大学医学院免疫学研究所和兄弟院校的第一线教师根据多年的教学和科研实践，充分考虑基因工程学及分子生物学在21世纪生命科学中的重要地位和新时期医学院校培养目标，坚持理论联系实际的原则编写而成。本书围绕基因工程学的基本理论、基本技术及基因工程学在医学中的应用，系统阐述了基因克隆、基因表达及多种常用基因工程学技术的原理及方法。具有简明扼要、可操作性强等特点。本书适合医学院7年制、8年制学生，临床和基础各科研究生，以及综合大学相关专业研究生，也可作为本科生选修课及青年医师进修、自学用书。

图书在版编目(CIP)数据

基因工程学/孙汶生等主编 .—北京:科学出版社,2004.8
中国科学院教材建设专家委员会规划教材
ISBN 7-03-013199-1

Ⅰ.基… Ⅱ.①孙… ②曹… ③马… Ⅲ.基因-遗传工程-医学院校-教材 Ⅳ.Q78

中国版本图书馆CIP数据核字(2004)第037982号

责任编辑:夏 宇 吴茵杰 / 责任校对:包志虹
责任印制:刘士平 / 封面设计:卢秋红

科学出版社 出版
北京东黄城根北街16号
邮政编码:100717
http://www.sciencep.com
新蕾印刷厂 印刷
科学出版社发行 各地新华书店经销
*
2004年8月第 一 版 开本:787×1092 1/16
2004年8月第一次印刷 印张:21 1/2
印数:1—4 000 字数:500 000

定价:32.00元
(如有印装质量问题,我社负责调换〈环伟〉)

前　言

基因工程技术自20世纪70年代诞生以来，发展异常迅速，从简单的基因体外"拼接"，逐步形成了关于基因结构、功能研究的一整套理论和方法。

目前，基因工程技术已广泛应用于人类基因组学、生物学、遗传学、肿瘤学、传染病学等与基础医学和临床医学有关的各个研究领域，并大大加快了这些领域的发展，作为21世纪的医学生和医学工作者应该了解、掌握和运用基因工程的相关理论和技术。为此，我们感到应该为他们提供一本简明扼要介绍该领域基本理论和相关技术的教材或参考书。

山东大学医学院免疫学研究所自20世纪80年代开始为研究生开设基因工程学课程，致力于开拓学生科研思路，培养研究生科学思维和训练学生实验技能，深受广大学生的好评。此次，我们在总结20年教学经验的基础上，吸收基因工程学最新研究成果和技术，编写此书。本书是山东大学医学院免疫研究所全体教师和研究生共同努力的成果，其主要使用对象是医学院临床和基础各科研究生，以及综合大学相关专业研究生，也可作为本科生选修课及青年医师进修、自学用书。

基因工程学涉及遗传学、生物化学、分子生物学及微生物学，属于多学科交叉的边缘学科。为了避免与其他学科的重复，本书在编写过程中以基因工程技术为主线，围绕基因克隆、体外重组、基因表达及基因工程学在医学中的应用，重点介绍相关技术的原理、技术和最新发展，突出实用性，注重理论与实验相结合。

全书共分为四篇。第一篇包括第一至第三章，概括介绍基因工程学发展动态及相关基本理论，包括DNA大分子的结构、特性、功能及基因工程的基本过程，力图对基因工程学的基本内容和新进展作一个框架式的介绍，其内容也涉及了当前许多崭新的研究领域，如基因组学、功能基因组学等。第二篇基因工程学基本理论与技术，包括第四章至第十二章，主要介绍基因工程常用技术及相关理论，通过这部分的讲述，读者可以了解有关基因克隆、基因表达的基本策略和过程，以及基因工程学中常用的工具酶、载体的概念、种类和当前基因工程中广泛应用的PCR、测序及芯片技术。第三篇基因工程在医学中的应用，包括第十三章至第十九章，主要介绍目前基因工程学在医学中应用后产生的新技术、新发展，如基因工程抗体、基因表达调控研究、功能基因筛选、基因治疗及利用基因工程技术制备新型实验动物模型等。第四篇是基因工程学实验技术，该

部分内容主要是结合前面介绍的理论,以基因克隆、基因表达为主线设计实验,力图通过这些实验使读者掌握基因工程的基本技术和技能。书后编写了七项附录,提供了一些常用培养基、试剂的配制、基因工程常用数据等,供研究者方便应用。为帮助读者能够更好地阅读、理解相关外文原著,本书收录了基因工程常用词汇及名词解释以供参考。本书编写中,山东省医学科学院韩金祥院长给予了大力支持并撰写"生物芯片的技术原理与应用"一章,泰山医学院魏然教授也为本书的出版给予了大力支持,在此一并表示致谢。

基因工程学是一门飞速发展的交叉学科,本书力图做到实用性和准确性,并尽量将基因工程学的最新成果纳入编写内容,然而,由于学科发展迅猛、编写时间仓促及编者水平所限,可能在内容、文字、编排及图表方面存在疏漏和不足之处,恳切希望读者和同道们指正。

孙汶生

2004 年 3 月于济南

目　　录

第一篇　概　　论

第二篇　基因工程学的基本理论与技术

第三篇 基因工程学的应用

第四篇 基因工程实验技术

第一篇　概　论

Part Ⅰ. Conspectus

第一章　基因与基因组

Chapter 1. Gene and Genome

第一节　基因的概念与特性

基因(gene)最初是遗传学概念，随着生物学科的不断发展及生命科学理论与技术研究的不断深入与完善，人们对基因本质的认识及其定义也在不断深化。

一、基因的概念与分子生物学定义

1926 年，Morgan 从遗传学的角度提出基因是位于染色体上的基本遗传单位，基因既是携带遗传信息的结构单位，又是控制遗传性状的功能单位。

早期研究曾认为遗传物质是蛋白质，直至 1944 年 Avery 等从肺炎链球菌转化实验的研究首次证明，控制某些遗传特性的物质是 DNA，而不是蛋白质(见图 1-1)。

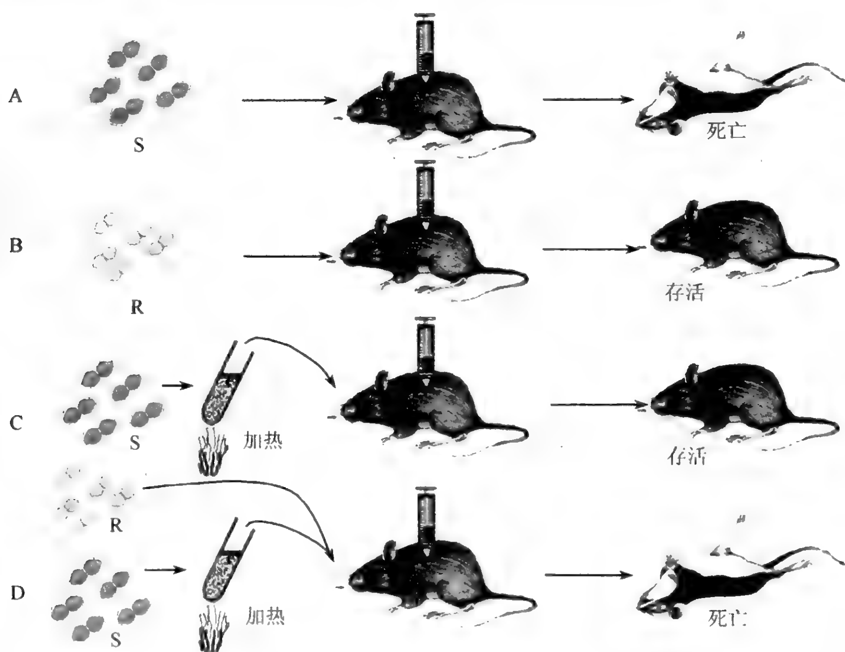


图 1-1　Avery 的肺炎链球菌转化实验(1944 年)

A. 注射有荚膜(S 型)的致病性肺炎链球菌，小鼠死亡　B. 注射突变的(R 型)非致病性肺炎链球菌，小鼠存活　C. 注射加热杀死的 S 型，小鼠存活　D. 注射活的 R 型与加热杀死的 S 型的混合物，小鼠死亡

1953 年 Watson 和 Crick 提出了 DNA 分子的双螺旋结构模型，阐明了 DNA 自我复制的机制，推测 DNA 分子中的碱基序列储存了遗传信息。1961 年法国科学家 F. Jacob 和 J. Monod 以及其他科学家相继发表了他们对调控基因的研究，证实了信使 RNA(mRNA)携带着从 DNA 到蛋白质合成所需要的信息；发现了遗传密码，提出了遗传信息储存于核酸之中；DNA 通过转录和翻译控制蛋白质合成，从而将 DNA 双螺旋结构与 DNA 功能联系起来。

事实上，并不是所有的基因都是 DNA。有些基因还包括不同类型的 RNA，如 tRNA、rRNA 等。编码蛋白质或 RNA 的基因均称为结构基因。此外，DNA 还含有只起调节功能的片段或序列，称为调节序列。调节序列可提供信号，表示结构基因的开始和结束，或参与启动或关闭结构基因的转录，或作为复制起始点起作用。

因此，基因的分子生物学定义是：基因是编码功能性蛋白质多肽链或 RNA 所必需的全部核酸序列(通常是 DNA 序列)，负载特定的遗传信息并在一定条件下调节、表达遗传信息，指令蛋白质合成。根据此定义，一个基因应包含了编码蛋白质肽链或 RNA 的序列、为保证转录所必需的调控序列、内含子以及相应编码区上游 5′端和下游 3′端的非编码序列。

二、基因的一般特性

基因有三个基本特性：①基因可自我复制，通过 DNA 半保留复制，保证物种的遗传稳定性；②基因决定生物表型或性状，即基因通过转录和翻译决定多肽的氨基酸序列，由此确定某种酶或蛋白质的性质，表达某一特定的生物性状；③很多因素可导致基因突变，新突变的等位基因一旦形成，就可通过自体复制，在其后的细胞分裂中保留下来传给子代，产生新的生物性状。

基因按其功能主要分为结构基因和调控基因：

1. 结构基因　结构基因(struicture gene)是能决定某些多肽链(蛋白质)或酶分子结构的基因。结构基因的突变可导致特定蛋白质(或酶)一级结构的改变。

2. 调控基因　调控基因(regulator gene)是具有调节控制结构基因表达功能的基因。调控基因的突变可以影响一个或多个结构基因的功能，导致蛋白质(或酶)量或活性的改变。

此外，有的基因只转录而不翻译，如核糖体 RNA(ribosomal RNA, rRNA)基因，其功能仅是转录 rRNA；转运 RNA(transfer RNA, tRNA)基因，其功能仅是转录 tRNA。

第二节　DNA 的结构与性能

核酸(nucleic acid)是生物的遗传物质基础，也是基因的基本结构。核酸的化学组成、分子结构符合遗传物质的稳定性、连续性及多样性的要求。

一、核酸的化学组成

核酸结构的基本单位是核苷酸(nucleotide acid)，每个核苷酸由 1 个磷酸、1 个五碳糖和 1 个碱基 3 部分组成。核酸有两类：一类是脱氧核糖核酸(deoxyribonucleic acid, DNA)，

DNA 中的脱氧核糖核苷酸主要由 4 种碱基，即腺嘌呤(adenine, A)、鸟嘌呤(guanine, G)、胞嘧啶(cytosine, C)和胸腺嘧啶(thymine, T)及脱氧核糖(deoxyribose)和磷酸构成；另一类是核糖核酸(ribonucleic acid, RNA)，RNA 分子中的核糖核苷酸主要由碱基 A、G、C 和尿嘧啶(uracil, U)及核糖(ribose)和磷酸构成(见图 1-2)。

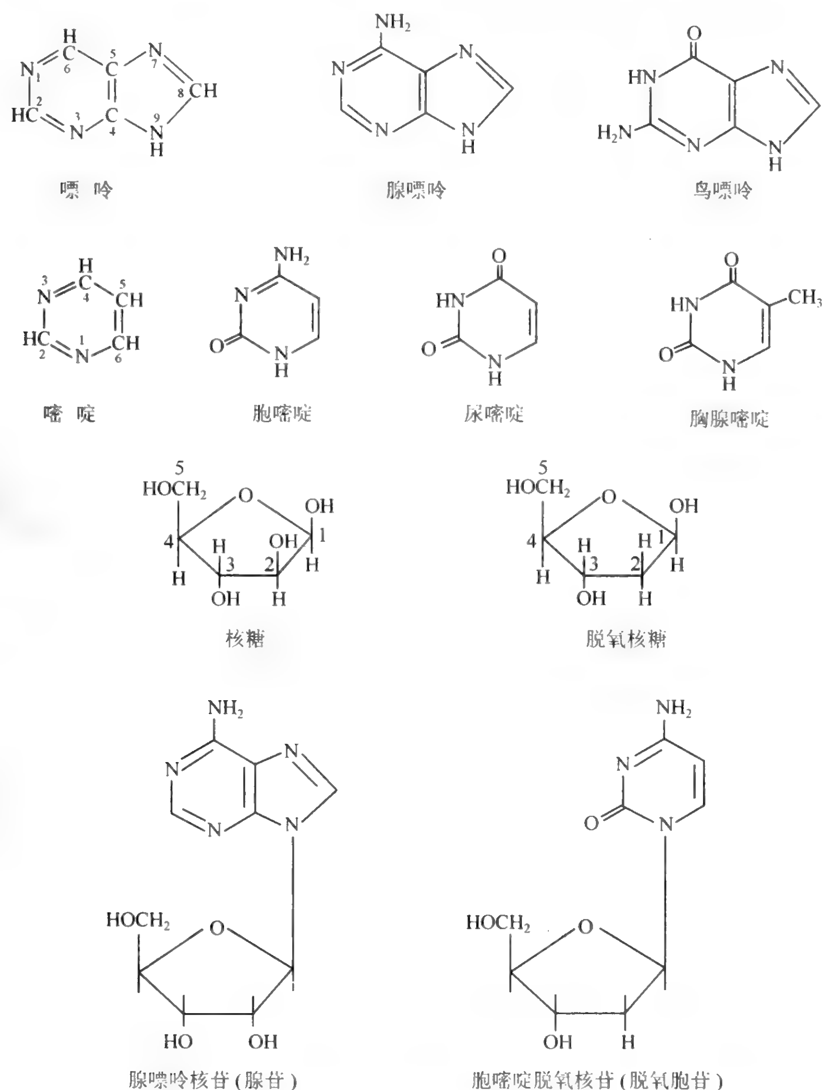


图 1-2　核酸的化学组成

DNA 分子中含有磷，故常用放射性核素 ^{32}P 做 DNA 标记。

二、DNA 的分子结构

（一）DNA 的一级结构

DNA 的一级结构是指 4 种核苷酸的连接及其排列顺序。DNA 分子是 4 种脱氧核苷酸经 3′-5′磷酸二酯键聚合而成，故也称多核苷酸(polynucleotide)。

（二）DNA 的二级结构

1953 年，Watson 和 Crick 提出了 DNA 分子双螺旋结构模型，其要点是：DNA 分子是由 2 条平行的多核苷酸链围绕同一中心轴构成的右手双螺旋结构。多核苷酸的方向由核苷酸间的磷酸二酯键的走向决定，一条从 5′→3′，另一条从 3′→5′，两条链呈反向平行排列(antiparallel)，彼此由氢键相连，G 与 C 配对(G≡C)，A 与 T 配对(A＝T)，形成 DNA 的二级结构。图 1-3 表明 DNA 的分子骨架。

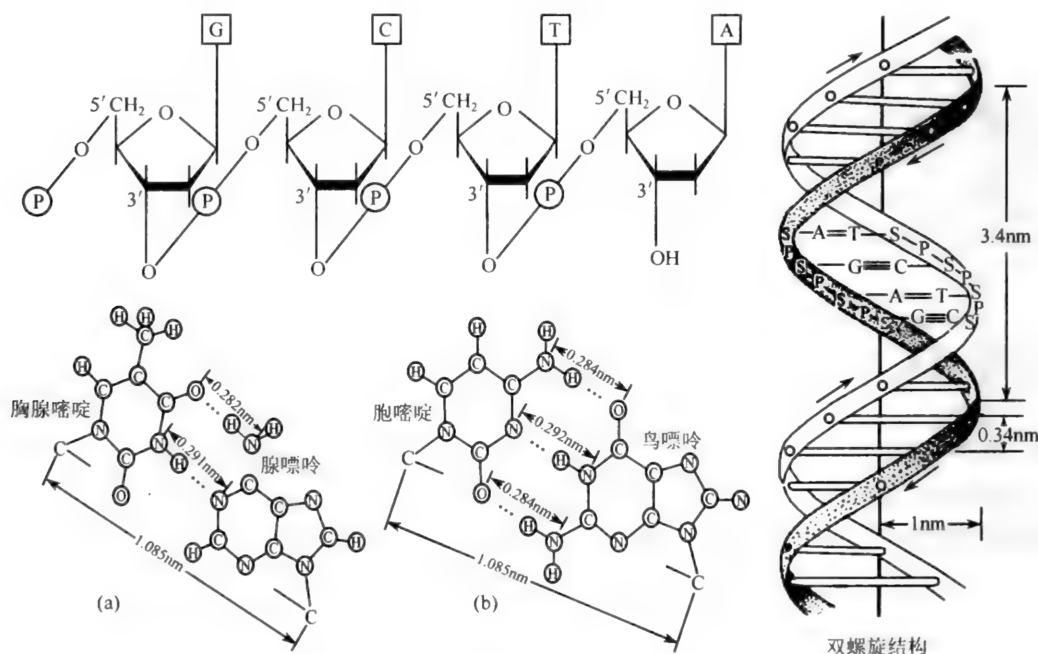


图 1-3　DNA 双螺旋结构及碱基配对示意图

（三）DNA 的超螺旋结构

生物体内绝大多数的 DNA 是以超螺旋(supercoil, SC)的形式存在。细菌及质粒 DNA 均为共价闭环(covalent close circle, CCC)双链 DNA，经 α 螺旋卷曲为超螺旋结构。当超螺旋结构的 DNA 任一单股链出现缺口时，螺旋则松开呈环状，称为开环(open circle, OC)；环状 DNA 分子不稳定，易被限制性内切酶切断，形成线型(linear, LC)DNA。由于超螺旋

DNA 有较大的密度，在凝胶电泳时涌动速度较快，在离心场中的移动速度也较线性或环状的 DNA 快，故应用凝胶电泳或超速离心法可以很容易的将超螺旋 DNA 分开(图 1-4)。

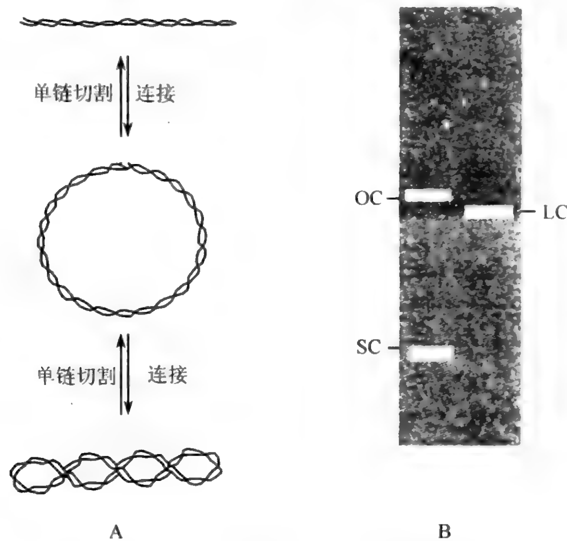


图 1-4　DNA 的构型(A)与电泳位置(B)

由于相对分子质量相同而构型不同的 DNA 在电场中泳动速度不同，所以 DNA 相对分子质量测定时均需采用线型 DNA。

三、DNA 的功能与性质

1. DNA 的吸收光谱　DNA 分子的吸收光谱高峰为 260nm，紫外线照射可引起 DNA 损伤断裂。

2. DNA 分子在电场中的迁移　DNA 分子在碱性条件下带负电，凝胶电泳中向正极泳动。分子大小及构象不同的 DNA 分子其迁移率不同，从而可以分开。溴化乙啶(EB)可嵌入 DNA 双链的配对碱基之间，在紫外线激发下呈现橙红色荧光，可用于 DNA 分析及相对分子质量测定。

3. DNA 的变性　在一定物理或化学因素的作用下，DNA 的双链间的氢键断裂，形成单链的过程，叫做 DNA 的变性(denature)。加热到一定温度，DNA 双链分开，称热变性。DNA 对波长为 260nm 的紫外线具有吸收作用，且单链 DNA 对紫外线的吸收值高于双链 DNA。由于 DNA 变性而使其溶液对紫外线的吸收值增加的现象称为增色效应，常以此特性进行 DNA 观察和测定。使一半的螺旋结构解体时的温度被定义为解链温度(temperature melting，T_m)。T_m 对 DNA 的 PCR 扩增十分重要(见第九章)。

4. DNA 的复性　将 DNA 的变性因素去除后，两条单链根据碱基互补原则，恢复成双链的过程，称为 DNA 的复性(renature)。热变性后慢慢冷却，两条单链恢复成双链螺旋结构的过程，称为退火(annealing)。来源不同的两条核酸单链，根据碱基互补的原则形成双链的

过程，称为杂交(hybridization)。杂交可发生在 DNA/DNA 之间，也可发生在 DNA/RNA 之间。核酸杂交的种类很多，如 Southern blot、Northern blot、原位杂交(*in situ* hybridization)等。采用标记 DNA 探针进行杂交，可以检测特定 DNA 的存在情况(见第十章)。

第三节 DNA 的复制与表达

一、DNA 复制

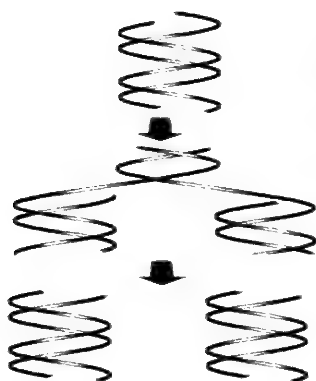


图 1-5 双螺旋 DNA 半保留复制过程

DNA 复制过程中，首先碱基间氢键断裂，双螺旋解旋并松开，然后每条多核苷酸链各以自己为模板(template)吸收游离核苷酸，按碱基互补原则，进行氢键结合。在聚合酶作用下，合成新的互补链，与原来的模板单链并列盘旋，形成稳定的双螺旋结构(图 1-3)。由此新形成的 2 个 DNA 分子与原来 DNA 分子的碱基序列完全一样。每个子代 DNA 分子的一条链来自亲代 DNA，另一条链则是新合成的，所以这种复制方式称为半保留复制(semi-conservative replication，图 1-5)。

二、基因表达

基因表达(gene expression)是指细胞在生命过程中，把储存在 DNA 序列中的遗传信息经过转录和翻译，转变成具有生物活性的蛋白质分子。生物体内的各种功能蛋白质和酶都是由相应的结构基因编码的(图 1-6)。

(一) 转录过程

在 RNA 聚合酶的催化下，以 DNA 为模板合成 mRNA 的过程称为转录(transcription)。在双链 DNA 中，作为转录模板的链称为模板链(template strand)，或反义链(antisense strand)；另一条链称为编码链(coding strand)，或有义链(sense strand)。编码链与其转录产物的差异仅在于 DNA 中 T 变为 RNA 中的 U。在含许多基因的 DNA 双链中，每个基因的模板链并不总是在同一条链上，亦即一条链可作为某些基因的模板链，也可是另外一些基因的编码链。

转录后要进行加工，转录后的加工包括：

1. 剪接　一个基因的外显子和内含子都转录在一条原始转录物 RNA 分子中，称为前 mRNA(pre-mRNA)，又称不均一核 RNA(heterogeneous nuclear RNA，hnRNA)。前 mRNA 分子既有外显子序列又有内含子序列，还包括编码区上、下游的非翻译序列。必须除去内含子序列，才能产生成熟的、有功能的 mRNA 分子，这个过程称为 RNA 剪接(RNA splicing)。剪切发生在外显子的 3′末端的 GT 和内含子 3′ 末端与下一个外显子交界的 AG 处。

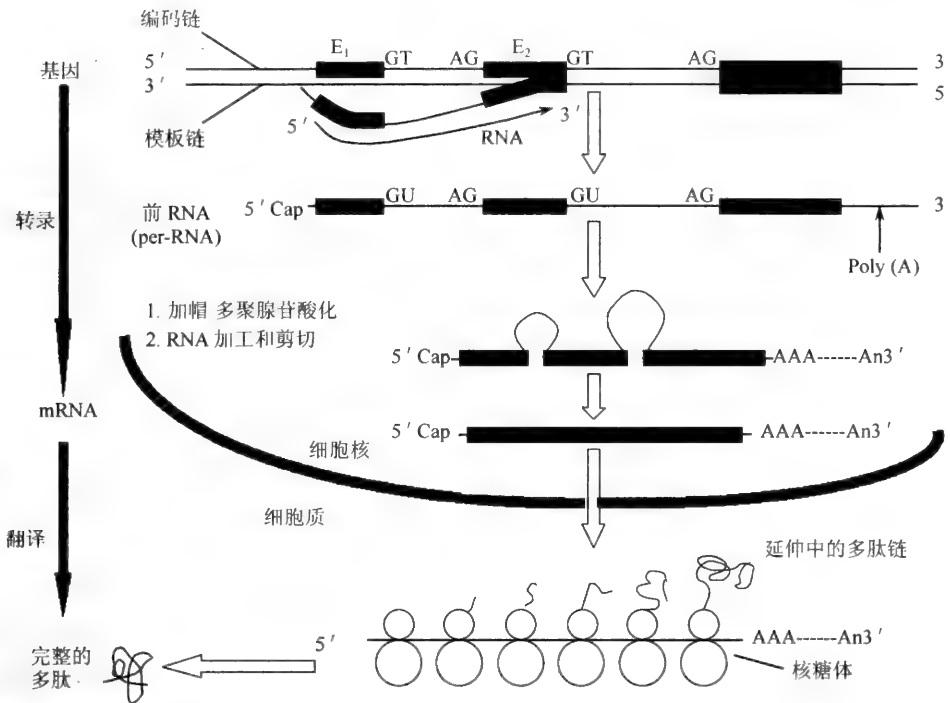


图 1-6　真核生物结构基因表达(DNA→RNA→蛋白质)流程图

2. 加帽　几乎全部的真核 mRNA 5′端都具“帽子(cap)”结构。虽然真核生物的mRNA的转录以嘌呤核苷酸三磷酸(pppA 或 pppG)领头,但在 5′端的一个核苷酸总是 7-甲基鸟核苷三磷酸(m7GpppAGpNp)。mRNA 5′端的这种结构称为帽子。不同真核生物的 mRNA 具有不同的帽子。

mRNA 帽子结构的功能:①能被核糖体小亚基识别,促使 mRNA 和核糖体的结合;②m7Gppp结构能有效地封闭 RNA 5′末端,以保护 mRNA 免受 5′核酸外切酶的降解,增强 mRNA 的稳定性。

3. 加尾　大多数真核生物的 mRNA 3′末端都有由 100～200 个腺嘌呤(A)组成的 polyA 尾巴。polyA 尾不是由 DNA 编码的,而是转录后的前 mRNA 以 ATP 为前体,由 RNA 末端腺苷酸转移酶,即 polyA 聚合酶催化聚合到 3′末端。

mRNA poly(A)尾的功能是:①有助于 mRNA 从细胞核到细胞质的转运;②避免在细胞中受到核酶降解,增强 mRNA 的稳定性。

(二) 翻译过程

真核细胞的转录以及加工都在细胞核内进行,但翻译过程则在细胞质中进行。

以 mRNA 作为模板,tRNA 作为运载工具,在有关酶、辅助因子和能量的作用下,将活化的氨基酸在核糖体(亦称核蛋白体)上装配为蛋白质多肽链的过程,称为翻译(translation),这一过程大致可分为肽链的起始、延长与终止 3 个阶段(图 1-7)。

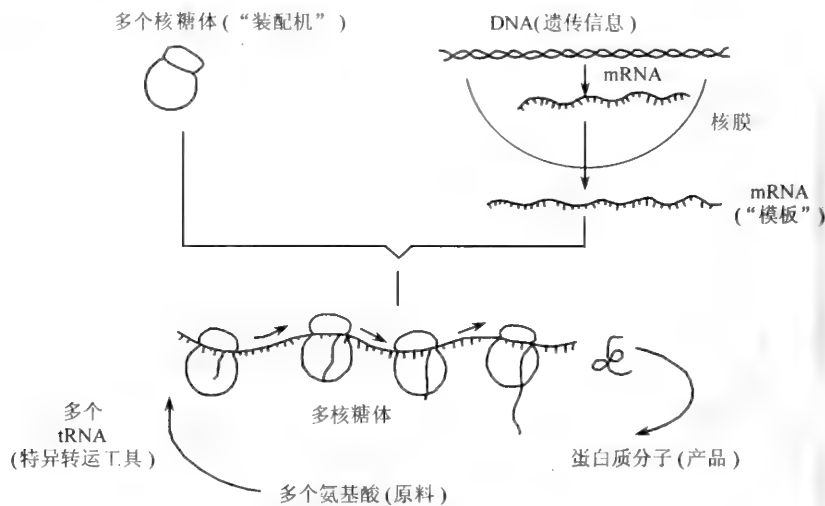


图 1-7 真核细胞的翻译过程

从核糖体上释放出来的多肽需要进一步加工修饰才能形成具有生物活性的蛋白质，这一过程称为翻译后加工(posttranslational processing)。翻译后的肽链加工包括肽链切断，某些氨基酸的羟基化、磷酸化、乙酰化、糖基化等。真核生物在新生多肽链翻译后将蛋氨酸裂解掉。

第四节 基 因 组

基因组(genome)是细胞或生物体的全套遗传物质，载有遗传信息的 DNA 大分子通常组装成染色体(原核生物中为染色质)，每条染色体存在成千上万个不同的基因。

人们曾经认为每个基因组的 DNA 量是不变的(包括数目、位置等)，但至 20 世纪 60 年代，在大肠杆菌的半乳糖操纵子的研究中发现了插入序列以及可转移位置的遗传因子后，才认识到基因组的某些成分的位置并非一成不变。尽管如此，各类生物的基因组仍有其基本的结构特点，物种不同，染色体数目不同，其基因数量不等，少至几个，多则 5 万～10 万个。不同物种的正常染色体数目见表 1-1。

表 1-1 不同生物的正常染色体数目

物种	染色体数目	DNA 含量/bp
λ噬菌体(bacterial phage λ)	1	5×10^4
大肠杆菌(*E. coli*)	1	4.2×10^6
啤酒酵母(yeast)	34	1.4×10^7
果蝇(fruit fly)	8	1.4×10^8
小鼠(mouse)	40	4.7×10^9
蛙(frog)	26	4.5×10^9
人(human)	46	3.2×10^9
鸡(chicken)	78	2.1×10^9

一、病毒基因组

病毒不能单独繁殖,只能在宿主细胞内进行复制。完整的病毒颗粒是由核酸和蛋白质组成的;其核心为核酸,构成病毒的基因组。病毒基因组的特点使其可以通过人工改造成为基因工程中良好的载体。

(一) 病毒基因组的结构

病毒的基因组由 DNA 或 RNA 组成,每种病毒颗粒只含有一种核酸。病毒基因组的 DNA 或 RNA 可以是单链的(single strand, SS),也可以是双链的(double strand, DS);可以是环状分子,也可以是线性分子。多数 DNA 病毒的基因组是双链 DNA 分子,如乙型肝炎病毒的基因组是双链环状 DNA,腺病毒、疱疹病毒的基因组为双链线状 DNA;多数 RNA 病毒的基因组是单链 RNA 分子,如脊髓灰质炎病毒。

(二) 病毒基因组的特点

1. 病毒基因组小,基因数少。DNA 病毒基因组的大小差别很大。如乙型肝炎病毒只有 3.2kb,只能编码几种蛋白质;而痘病毒基因组有 300kb,可以编码几百种蛋白质。

2. 带有重叠基因,即同一段 DNA 片段可以编码 2～3 种蛋白质分子。这种结构可使较小的基因组携带更多的遗传信息。例如乙型肝炎病毒(HBV)(图 1-8)。虽然有些基因完全重叠在一起,但它们的开放读码框架(open reading frame, ORF)不同,因此翻译的氨基酸序列及表达的蛋白质完全不同。

3. 噬菌体的基因是连续的,基因组中无内含子。感染真核细胞的病毒基因是不连续的,具有内含子。

4. 病毒基因组大部分为编码蛋白质的结构基因,一小部分序列为基因之间的间隔区和基因表达的调控序列。

5. 病毒基因组 DNA 序列中功能相关的蛋白质基因或 rRNA 基因常集中在基因组的一个或几个特定的部位,形成一个功能单位或转录单元。

6. 除反转录病毒基因组有两个拷贝外,其他病毒基因组中每个基因只有一个拷贝。

二、原核生物基因组

原核生物的核物质分散在细胞质中,无核膜包围,只有染色质,不形成染色体。通常多数为双链环状结构,少数为线状。由于其核物质只有 DNA 分子,不含组蛋白,没有核仁,不形成明显的细胞核,无核膜,故称为类核(nucleoid)。例如大肠杆菌染色质是由 4.2×10^6bp 组成的双链环状 DNA 分子,约有 3 000～4 000 个基因,目前已经定位的基因已达 900 多个。

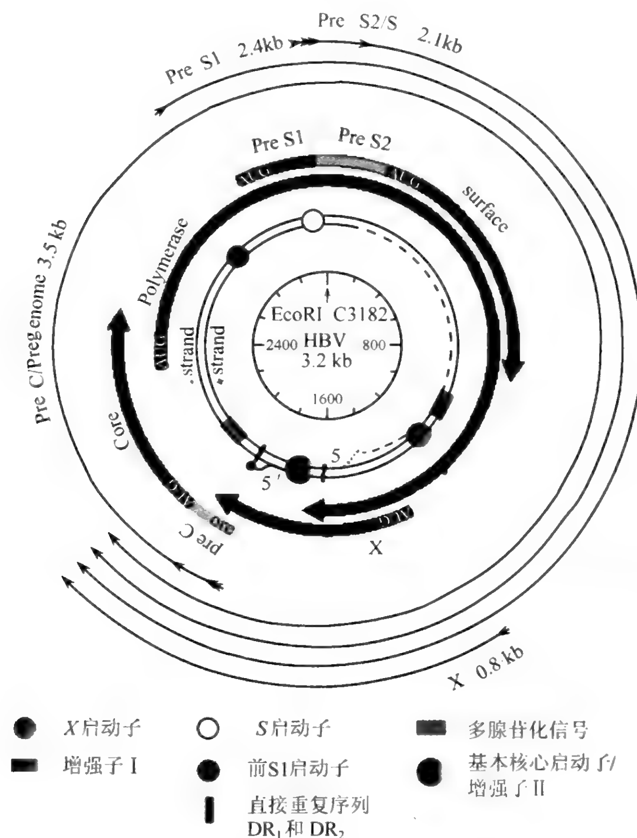


图 1-8 HBV 基因组结构

(一) 细菌染色质基因组的结构特点

1. 细菌染色质基因由环状双链 DNA 组成,经高度折叠、聚集,形成致密的类核。类核中央部分由 RNA 和支架蛋白组成,外围是双链闭环的 DNA 超螺旋。DNA 链上与复制、转录有关的信号区域优先与细胞膜结合,这种连接有助于细胞膜对染色质的固定及细胞分裂时染色质的均匀分配。

2. 功能上相关的几个结构基因往往串联排列在一起组成操纵子(operon)结构,受上游调控区控制。转录时,几个基因转录在一条 mRNA 链上,再分别翻译成各自的蛋白质。

3. 结构基因序列是连续排列的,无内含子,在转录后不需要加工修饰。mRNA 在 DNA 链上合成时,核糖体就开始将 mRNA 翻译成蛋白质,即边转录边翻译,翻译与转录偶联进行。

4. 细菌的 DNA 绝大部分都是编码蛋白质的功能基因。

5. 细菌基因组结构基因无重叠现象。

6. 细菌基因组中结构基因多为单拷贝,但编码 rRNA 的基因通常是多拷贝的,以适应细菌核糖体快速组装及蛋白质快速合成。

(二) 质粒

细菌除了在类核中含有较大环状染色质 DNA，胞质中还含有小的双链环状 DNA 分子结构，称为质粒(plasmid)。质粒是细菌的特殊结构，多在 $4\times10^6\sim100\times10^6$bp 范围内，具有良好的自我复制和调控系统。质粒携带的遗传信息通过复制，产生子代质粒，并在细胞分裂时进入子代细胞。质粒广泛存在于细菌，有些酵母菌和其他真菌中也发现有质粒存在。大肠杆菌质粒结构见图 1-9。

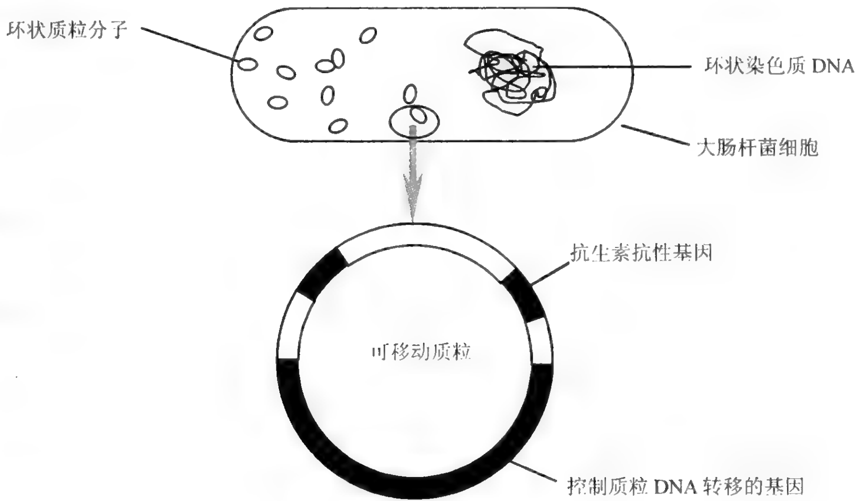


图 1-9　大肠杆菌质粒结构示意图

质粒并不是细菌生命活动所必需的，但其编码的性状大多数对细菌起保护作用，使细菌获得有利于其生存的特性。

质粒携带有许多基因，可以控制许多重要的生物学形状。例如携带 β-内酰胺酶的质粒使宿主菌对 β-内酰胺类抗生素如青霉素和氨苄西林产生耐受性。

质粒可以从一个细菌转移至另一个细菌，可以在同种属的细菌或在不同种属的细菌之间转移。携带的性状也随之转移。质粒可以从抗生素耐受细菌转入对抗生素敏感的细菌，使后者变为耐药菌。质粒在致病菌中播散，形成多重耐药菌株。某些质粒还能与染色质整合，其复制则受染色质的控制。

把所需目的基因插入到质粒中，并导入受体细胞，通过质粒的复制及转录，使受体细胞能够产生由外源基因编码的蛋白质。质粒的特性使其在基因工程中成为应用十分广泛的载体工具(见第五章)。

三、真核生物基因组

真核生物中 DNA 结构复杂，表达调控多样。DNA 在细胞中的包装是一种独特的、紧凑和高度凝聚的结构，经过多级螺旋盘绕压缩，与蛋白质结合形成染色体。整个基因组分布在细胞核内的多条染色体中，外面有核膜包裹。真核生物的染色体在细胞周期的大部分时间

内以染色质(chromatin)的形式存在,染色质的基本单位是核小体(nucleosome)。由于染色体 DNA 存在于细胞核中,故 mRNA 在细胞核内合成后,要经过修饰、加工,如剪切、加多聚腺苷酸(poly A)尾等,然后运输到细胞质,指导蛋白质合成。

真核细胞除细胞核 DNA 之外,细胞器、线粒体也含极少量的 DNA(不到细胞全部 DNA 的 1%)。但受精和分裂的卵细胞中线粒体较多,线粒体 DNA(mtDNA)的总量也相应的多一些。线粒体 DNA 分子甚小,全长约有 14kb,为环状双螺旋结构,编码线粒体 tRNA、rRNA 和少数几种线粒体蛋白质。95% 以上的线粒体蛋白质是由核 DNA 编码的。当细胞分裂时,线粒体 DNA 亦进行复制,并进入子代细胞器中。植物中存在叶绿体 DNA。

(一) 真核生物基因组的一般特点

1. 真核生物的基因组比较庞大,例如人的单倍体基因组由 3.2×10^9bp 组成,含有大约 3 万个基因。具有许多复制起始点,但每个复制子的长度较小。

2. 真核生物基因组 DNA 与蛋白质结合形成染色体,储存于细胞核内,除配子细胞外,体细胞内的基因是双倍体(diploid),即有两份同源的基因组。

3. 真核细胞基因无操纵子,转录产物为单顺反子,一个结构基因经过转录和翻译生成一个 mRNA 分子和一条多肽链。

4. 真核生物基因组存在大量重复序列,重复次数可达百万次以上。

5. 真核生物基因组中非编码区域多于编码区域。

6. 真核生物大部分基因含有内含子(间隔序列),因此结构基因是不连续的,称为断裂基因(split gene)。

7. 真核生物基因组中许多来源相同、结构相似、功能相关的基因在染色体上成串存在,这样的一组基因称为基因家族(gene family)。多基因家族是真核生物基因组织的一个重要特征。

8. 真核生物的基因大小差别很大,例如,人类血红蛋白的基因长仅约 1 700bp,而人类假肥大型营养不良症(duchenne muscular dystrophy, DMD)基因全长为 2 300kb。

(二) 真核生物基因组的 *C* 值矛盾

同一物种的基因组 DNA 含量总是恒定的,一个单倍体基因组中的全部 DNA 量称为该物种 DNA 的 *C* 值(*C* value)。一般来说,*C* 值随着生物的复杂性的增加而增加。结构和功能越复杂,需要的基因越多,其 *C* 值越大。但某些植物或两栖类,它们的 DNA 含量比人类高出几十倍,这种形态学复杂程度与 *C* 值大小不一致的反常现象称为 *C* 值矛盾(*C* value paradox),其机制和意义尚不明确。

(三) 真核生物基因结构

真核细胞核内基因的 DNA 序列由编码序列和非编码序列两部分构成。编码序列被非编码序列隔开,是不连续的。基因结构包括 4 个区域:①编码区,包括外显子与内含子;②前导区,位于编码区上游,相当于 RNA 5′末端非编码区(非翻译区);③尾部区,位于 RNA 3′编码区下游,相当于末端非编码区(非翻译区);④调控区,包括启动子和增强子等。基因编码区的两侧也称为侧翼序列。

1. 外显子和内含子　大多数真核生物的基因为不连续基因(interrupted or discontinuous gene)。所谓不连续基因就是基因的编码序列在DNA分子上是不连续的,被非编码序列所隔开。编码的序列称为外显子(exon),是基因编码表达多肽链的部分;非编码序列称为内含子(intron),又称插入序列(intervening sequence, IVS)。内含子可转录,但在前mRNA(pre-mNRA)时被剪切掉。内含子的核苷酸数量可比外显子多许多倍。如果一个基因有 n 个内含子,一般基因的外显子被分隔的数目为 $n+1$。外显子在mRNA转录过程中,受特定酶的作用,被拼接成具有表达活性的完整的结构基因。

2. 外显子-内含子接头　每个外显子和内含子接头区都有一段高度保守的一致序列(consensus sequence),即由内含子5′末端大多数是GT开始,3′末端大多是AG结束的序列,称为GT-AG法则,是普遍存在于真核基因中RNA剪接的识别信号。

3. 侧翼序列　在第一个外显子的上游和最末一个外显子的下游是一段不被翻译的非编码区,称为侧翼序列(flanking sequence)。侧翼序列含有基因调控序列,对该基因的活性有重要影响。

4. 启动子　启动子(promoter)具有促进启动转录过程的功能,与基因转录的起始、调节及转录激活功能密切相关(详见第七章)。

5. 增强子　增强子(enhancer)不能启动基因的转录,但有增强转录的作用。一般位于真核基因转录起始点的上游或下游。增强子序列可与特异性细胞因子结合从而促进转录的进行。研究表明,增强子通常有组织特异性,这是因为不同细胞核有不同的特异性细胞因子与增强子结合,从而对基因表达有组织、器官、时相与强度不同的调节作用。

例如人类单拷贝胰岛素基因5′末端上游约250 bp处有一组织特异性增强子,在胰岛β细胞中有一特异因子可作用于该区以增强胰岛素基因的转录和翻译;其他组织中无此因子。这正是胰岛素基因只有在胰岛β细胞中才得以很好表达的原因。

6. 终止子　终止子(terminator)具有转录终止的功能。终止子是在转录终止点之前的一段回文序列,约7～20bp。回文序列的对称轴一般距转录终止点16～24bp(图1-10)。

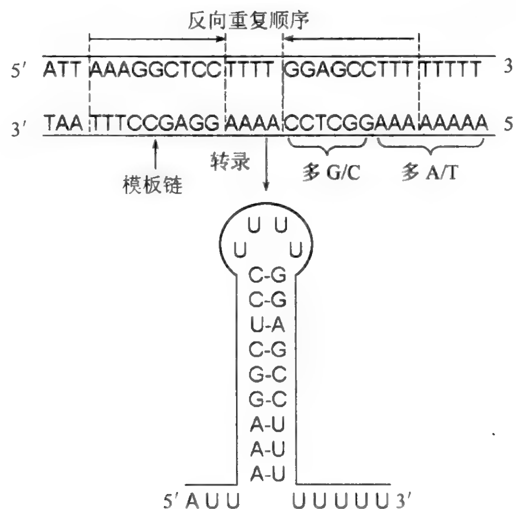


图1-10　转录终止子序列图

在回文序列的下游有 6～8 个 A—T 对，因此，这段终止子转录后形成的 RNA 具有发夹结构，并具有与 A 互补的一串 U，因为 A—U 之间氢键结合较弱，因而 RNA/DNA 杂交部分易于拆开，这样对转录物从 DNA 模板上释放出来是有利的，也可使 RNA 聚合酶从 DNA 上解离下来，实现转录的终止。

四、人类基因组

人类基因组(human genome)包括细胞核内的基因组及细胞质内线粒体基因组，其结构不同，各有特点。

(一) 细胞核基因组

人类细胞的全部遗传信息(基因)都编码在线状的 DNA 分子上。每个单倍体基因组约含 3.2×10^9bp。人类基因的平均长度为 1～1.5kb，所以基因组足以编码 1.5×10^6 种蛋白质，但实际上编码蛋白质的结构基因只不过 3 万个，仅占总基因组的 2%～3%。其余的 DNA 序列包括基因之间的间隔序列、基因内插入序列、重复序列等。研究证实，其中一些有着特殊的功能，包括调节基因的表达，增强同源染色体之间的配对和重组，维持染色体结构，调节前 mRNA 的加工以及参与 DNA 的复制等。但绝大多数重复序列的功能目前尚知之甚少。

1. 单一序列(单拷贝序列)　单一序列(unique sequence)又称非重复序列，约占基因组的 60%～65%，这种序列在一个基因组中一般仅有单个或几个拷贝，大多数编码蛋白质和酶基因属于此类。单一序列的两侧往往为散在分布的重复序列。因为大多数结构基因是单拷贝序列，所以结构基因的突变很容易造成遗传性状的改变或产生遗传性疾病。

2. 重复序列　重复序列(repetitive sequence)是指在一个基因组中具有很多拷贝的序列，又可分为几类：

(1) 高度重复序列(highly repetitive sequence)：高度重复序列一般不能转录，它们参与染色体结构的维持，形成结构基因间隔，可能与减数分裂时同源染色体的联会配对有关。其长度可能 2、4、6、8 等几个 bp，较长的序列可达 200bp，但其重复拷贝数可多达 10^6 次以上。例如端粒就是由高度重复序列的卫星 DNA(satellite DNA)构成。

(2) 中度重复序列(moderately repetitive sequence)：一般都是不编码的序列，其长度为 300～7 000bp。大多数与单拷贝基因间隔排列，少数在基因组中成串排列在一个区域。可重复数十至数万($<10^5$)次，在基因调控中起重要作用，包括开启或关闭基因、促进或终止转录、DNA 复制的起始并参与前 mRNA 加工等。

(3) 基因家族和基因簇：真核基因组中有许多来源相同、结构相似、功能相关的基因，这组基因称为基因家族(gene family)。基因家族的成员可以分布于几条不同染色体上，也可集中于一条染色体上。集中成簇的一组基因称为基因簇(gene cluster)。例如人类白细胞抗原(HLA)系统的 7 个连锁基因座位，排列成 *A-C-B-D-DR-DQ-DP*，形成一个基因簇。有些基因家族的成员并不集中排列为基因簇，而是散布在基因组中不同部位，如微管蛋白基因家族。

3. 假基因　在多基因家族中的某些成员并不产生有功能的基因产物，称为假基因(pseudogene)，如 Ψξ、Ψα、Ψβ 等。

此外，人类基因组中还有一些特殊的短序列位于各基因的侧翼，它们是起到调控作用的调节序列(启动子、增强子等)。有的与细胞恶性转化有关的原癌基因(proto-oncogene)等构成了多样性的人类基因组结构。

(二) 线粒体基因组

人类线粒体 DNA(mitochondrial DNA, mtDNA)是独立于细胞核染色体外的基因组，能自主复制，由 16 569 个碱基对组成，每一个 mtDNA 分子为环状双链 DNA 分子，外环为重链，内环为轻链(图 1-11)。基因组含有 37 个基因，其中 13 个为蛋白质基因，主要编码与细胞呼吸链代谢相关的蛋白质或酶。

图 1-11　人类线粒体基因组

H:重链；L:轻链；ND1—ND6:基因编码 NADH 脱氢酶亚单位；CO1—CO3:基因编码细胞色素 *c* 氧化酶亚单位；1—3CYB:基因编码细胞色素 *b*

人类线粒体基因组具有下列特点：

1. 人类线粒体的基因排列的非常紧凑，除与 mtDNA 复制及转录有关的一小段区域外，无内含子序列。基因之间有时有重叠，因此，mtDNA 的任何突变都会累及到基因组中一个

重要功能区域。

2. mtDNA 为高效利用的 DNA。有 5 个阅读框架,缺少终止密码子,仅以 U 或 UA 结尾。

3. mtDNA 的突变率高于核中 DNA,并且缺乏修复能力。

4. mtDNA 为母系遗传。

5. 部分 mtDNA 的密码子不同于核内 DNA 的密码子。细胞核内所列的密码是一种通用密码,但是真核生物线粒体的密码却有若干处不同于通用密码。如人类线粒体 UGA 不是终止密码子,而是色氨酸的密码子;AGA、AGG 是终止密码子,而不是精氨酸的密码子。这样,加上通用密码中的 UAA 和 UAG,线粒体共有 4 个终止密码子。

线粒体的密码表排列得比较整齐,各种氨基酸的密码子以及起始和终止密码子的数目或是 2,或是 4,或是 6。

近年发现,有些遗传病如 Leber 遗传性视神经病、肌阵挛性癫痫等与线粒体基因突变有关,因而线粒体基因结构已引起普遍关注。

(刘玉刚　孙汶生)

第二章　基因工程的基本概念、过程及研究策略

Chapter 2. Basic Concept, Process and Research Strategy of Genetic Engineering

第一节　基因工程的基本概念

基因工程是生物工程体系中的重要组成部分。现代生物工程体系的组成包括：基因工程、细胞工程、酶工程、蛋白质工程、微生物发酵工程、生化工程、生物传感器等。

基因工程(genetic engineering)，又称重组 DNA 技术(recombinant DNA technology)，是指在基因水平上，根据人们的需要以人工的方法取得供体 DNA 上某个或某些有用的基因，在体外重组于载体(质粒、噬菌体、病毒等)DNA 分子上，然后将重组 DNA 转入受体细胞进行无性繁殖(称之为"克隆")和行使正常功能(称之为"表达")，从而产生人类所需要的产物或创造新的生物类型的生物学技术。

基因工程的核心是 DNA 分子的体外重组技术。外源 DNA 插入载体分子所形成的杂合分子又称为嵌合 DNA 或 DNA 嵌合体(DNA chimera)。经重组技术而获得的具有新功能的微生物称"工程菌"。基因工程通过基因重组，使一个独特的 DNA 顺序在细胞内复制、扩增，该过程也称为分子克隆。

克隆(clone)源于希腊文"klone"，原意是用"嫩枝"或"插条"繁殖。在基因工程中克隆是指生物体通过细胞无性繁殖形成的基因型完全相同的后代个体组成的群体，简称为"无性繁殖"。克隆技术(cloning)则指通过 DNA 重组和克隆扩增后从众多的基因或细胞群体中获得目的基因或细胞的技术操作。该技术反映了细胞核分化技术、细胞培养和控制技术的发展，是人类在生物科学领域取得的一项重大技术突破。

克隆可根据其研究或操作的对象分为基因克隆、细胞克隆和个体克隆三大类。基因克隆是指在分子(DNA)水平上开展研究，以获得大量的相同基因及其表达产物；细胞克隆则是在细胞水平上开展研究，以获得大量相同的细胞；个体克隆则是经过一系列的操作产生一个或多个与亲代完全相同的个体，如克隆羊等。

基因工程按其发展和认识过程经历了以下几种术语的变化，即：①基因操作(gene manipulation)；②基因重组(gene recombination)；③基因工程(genetic engineering)；④分子克隆(molecular cloning)；⑤DNA 工程(DNA engineering)；⑥基因克隆(gene cloning)。

第二节 基因工程的基本过程及研究策略

基因工程技术的基本过程可概括为六大步骤(图 2-1)。

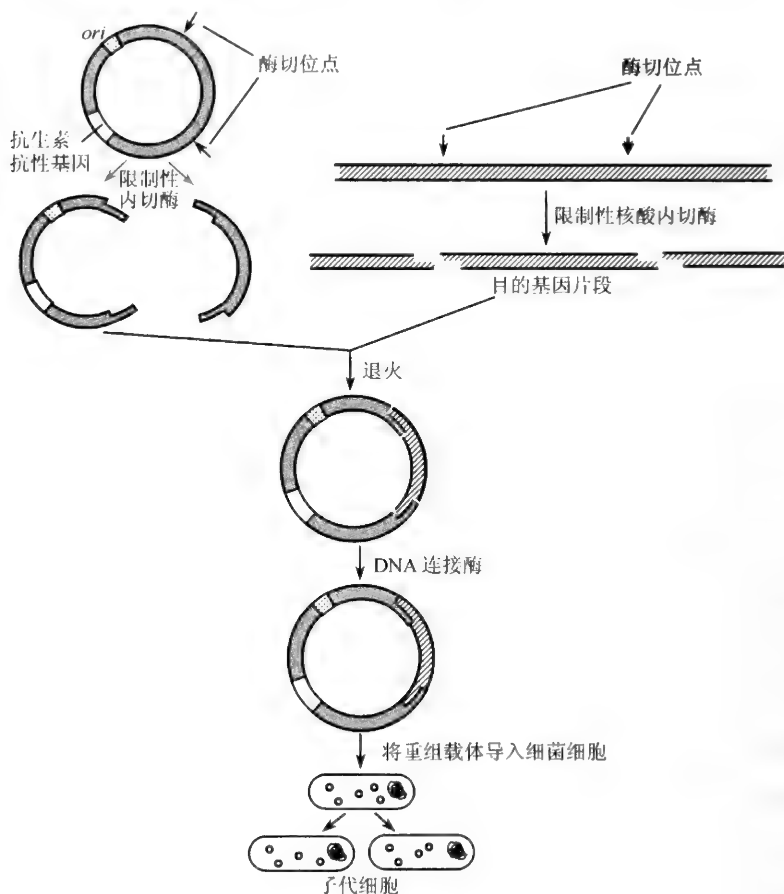


图 2-1 基因工程基本过程示意图

1. 人工合成或从生物基因组中,获得带有目的基因的 DNA 片段(donor DNA);

2. 选择或改造作为载体的 DNA(vector DNA);

3. 将带有目的基因的外源 DNA 片段连接到能够自我复制的并具有选择标记的载体分子上,形成重组 DNA 分子,即 DNA 重组(DNA recombination);

4. 将重组 DNA 分子引入受体细胞即转化(transformation)或转染(transfection);

5. 从大量的细胞繁殖群体中,筛选出获得了重组 DNA 分子的受体细胞(工程菌)并扩增,进一步抽提、纯化扩增的目的基因(clone、amplification);

6. 将目的基因克隆到表达载体上,导入受体细胞,使之在新的遗传背景下实现功能表达,产生出人类所需要的物质,即基因表达(gene expression)。

一、目的DNA的制备

(一) 直接分离DNA

适于遗传背景了解比较清楚的细菌染色体、质粒及病毒DNA的提取分离。其方法是将提取的DNA通过组建几种限制性内切酶(restriction enzyme, RE)的物理图谱进行基因定位,然后通过DNA的酶切片段电泳分离DNA,应用相应DNA探针确定目的DNA后,从制备电泳的凝胶中回收所需DNA。

(二) 鸟枪法克隆目的基因

在不了解DNA的遗传背景时,无法直接分离目的基因,可采用鸟枪法(shotgun)进行目的基因的克隆。

鸟枪法的基本技术是用限制性内切酶消化DNA,由此得到很多DNA的随机片段,将这些DNA片段与载体连接,组成由各种片段DNA构成的重组DNA,转化受体细胞,筛选含有插入DNA片段的重组阳性细胞。在一个适当的表达系统中,让所有重组体分别表达(建立基因库),通过探针杂交、表达产物分析或基因序列分析,选出正常表达目的基因的重组体,然后克隆扩增该重组体,酶切回收相应片段。

目前,以鸟枪法已建立了某些生物的基因库,为研究某些特定DNA片段提供了方便。

基因库(gene bank)是指某生物细胞内所存在的全套基因。基因库制备是将一种生物的全套遗传物质以酶消化后与载体DNA重组,转化受体细胞,并克隆扩增抽提制备DNA或采用物理方法(如超声波处理)制备DNA片段,使其收集到上百万的噬菌体颗粒中,经大量复制而形成了众多储有该细胞各个DNA片段的克隆,当制备的克隆数目多到可以把整个生物的全部遗传信息都包括在内时,这一组克隆的总体称为某一生物的基因库,也可称为基因文库(gene library)(见第六章)。

(三) 构建cDNA文库,筛选目的基因

真核细胞染色体DNA很大,特定基因只占染色体很小的部分(10^{-5}～10^{-7}单拷贝),如人染色体相对分子质量可达10^{12},又因真核细胞的基因内存在着间隔顺序(内含子),所以,不像原核细胞DNA分离那样易得。

基因工程中制备真核细胞基因一般是先分离纯化富含目的基因的mRNA,再反转录为互补DNA(complementary DNA,cDNA),然后进行cDNA的克隆扩增。包含有细胞mRNA信息的全部cDNA称为克隆cDNA文库,利用基因探针可由cDNA文库中钓取相应的目的基因。建立cDNA文库,筛选目的基因是真核细胞基因工程的基本策略。

近年,大规模cDNA测序和DNA芯片技术,为快速筛选目的基因或发现新基因提供了快速而准确的方法。

(四) 酶法或化学方法人工合成基因

1. DNA聚合酶链反应(polymerase chain reaction, PCR)　PCR技术可用于制备大量

DNA 特异片段。该法可从粗制品及降解的 DNA 模板中专一性大量扩增特异 DNA 片段，是 DNA“体外克隆”的一种形式(详见第九章)。

2. DNA 合成　利用 DNA 合成仪可合成任何已知序列的 DNA 片段或人为设计的 DNA 片段。目前，全自动合成仪化学合成 DNA 采用的是效率极高的亚磷酸三酯法(详见第六章)。

(五) mRNA 差异显示技术筛选差异表达基因

差异显示是利用相应的一对组织(如肝癌组织与正常肝组织)基因表达的差异，筛选出在某种组织中受某种因素调控的一个或一组差异表达基因的常用方法(详见第六章)。

(六) 差异蛋白质谱表达技术筛选功能基因

差异蛋白质谱表达技术是利用双向蛋白电泳对相关细胞中蛋白表达的差异进行直接比较，从而鉴定相关功能蛋白。经氨基酸序列分析，发现差异表达的相关功能基因(详见第六章)。

此外，重组 cDNA 表达文库的血清学分析(serological analysis of recombinant cDNA expression libraries, SEREX)是近年发展的一种功能基因的筛选方法。该法利用针对某抗原(如肿瘤抗原)的免疫血清筛选该抗原的 cDNA 表达文库中的阳性克隆，经进一步鉴定可获得编码该抗原的基因(见第十七章)。

二、载体 DNA 及其改造

(一) 基因工程中的载体

载体(vector)是指一个能够进行自我复制的复制子，它必须能携带外来 DNA 进入指定的受体细胞并在受体细胞内稳定保存、复制、扩增。

1. 常用的载体及其特点　基因工程中常用的载体包括：质粒(plasmid)、噬菌体(bacterial phage)、病毒(virus)、柯斯质粒(cosmid)等(详见第五、六章)。

常用载体的大小、结构、复制功能的差别很大，但其共同的特点为：

(1) 在宿主细胞中能独立的自我复制。

(2) 很容易从宿主细胞中分离、纯化。

(3) 载体 DNA 分子结构中有一段不影响其扩增的非必需区域，插入外源 DNA 片段后能被动的与载体一起复制、扩增。

2. 质粒　质粒是存在于细菌染色体之外的、能自我复制的双股闭环 DNA (dsDNA)，在基因工程中应用极广，根据在细胞内拷贝数的多少及特点，可将质粒分为两类：

(1) 严紧型质粒(stringent plasmid)：多为具有自身传递能力的大质粒，其 DNA 复制与宿主染色体 DNA 复制密切偶联，拷贝数低，约 1～5 个拷贝/细胞，如 pSC101。

所谓拷贝一般是指细菌内某一基因组或基因的复本；单拷贝(single copy)是指单倍体基因组中，仅出现一次的基因或 DNA 序列。

(2) 松弛型质粒(relaxed plasmid):多为相对分子质量小、不能自动传递的质粒,其复制在宿主细胞的松弛控制之下。拷贝数高,可达 10~200 个拷贝/细胞。

当氨基酸饥饿或加入蛋白合成抑制剂(如氯霉素)时,细胞染色体 DNA 和严紧型质粒 DNA 的复制停止,但松弛型质粒 DNA 能继续复制,甚至达到数千个拷贝。因此松弛性质粒被广泛用作基因工程中的载体。

(二) 作为载体 DNA 必须具备的条件

1. 载体 DNA 必须能被受体细胞所接受,进入宿主细胞后应能大量自我复制(克隆载体),作为表达载体还应具有促使外来 DNA 高效表达的调控区,使所携带的外来 DNA 能在其内忠实复制、稳定表达。

2. 有适应的限制性酶切位点,最好对不同的限制酶有各自的单一切点,能嵌入外来的 DNA。

3. 能赋予受体细胞易于检测的表型标志,以作为重组体的选择标记。

4. 载体 DNA 相对分子质量要相应的小,以保证外来 DNA 片段插入后构成重组体的稳定性。载体应易于获得、易于改造,具有多拷贝。

5. 携带外源 DNA 的幅度要宽,要非自我转移或在特定宿主细胞之外不能生存。

6. 从生物防护的角度应是安全的。

(三) 载体的改造

由于作为基因工程中的载体应具备以上几个条件,故自然界中的载体多需一定改造才能适于携带目的 DNA 进入受体细胞,易于筛选并能充分复制和表达。为此,常采用不同的策略进行载体改造,如在载体上引入一段具有多种限制性酶切位点的人工合成 DNA 序列,以便为外来 DNA 提供插入位点。这种序列通常称为多克隆位点(multiple cloning sites, MCS)序列。另外,也可在质粒中引进能够完善载体功能的辅助序列,如引入可供克隆筛选的含抗性基因的片段,或在真核表达载体上引入病毒启动子或增强子等元件。现以大肠杆菌质粒 pBR322 为例说明(图 2-2)。

pBR322 是由具有四环素耐药基因(tet^r)的 pMB$_1$ 与具有氨苄西林耐药基因(amp^r)的质粒 pBSF2134 共转大肠杆菌 RRI 受体细胞后经历一系列演化、重组、改造而产生的很有应用价值的载体颗粒(详见第五章)。

pBR322 相对分子质量为 2.6×10^6,长度 4.3kb,具有氨苄西林及四环素两个耐药基因(tet^r,amp^r),在 amp^r 上有限制性内切酶 *Pst* Ⅰ 的单一切点,在 tet^r 区域存在 *Hin*d Ⅲ,*Bam* H Ⅰ,*Sal* Ⅰ 的单一酶切点,当外源 DNA 整合到耐药基因特定位点时,可导致相应耐药基因失活,因此可依表型 Amp 及 Tet 作为正反选择标记(Tet^r,Amp^s 或 Tet^s,Amp^r)筛选重组体。

随着基因技术的发展,目前已开发出各种不同的、商品化的载体。无论从携带外源 DNA 的大小、特点,还是限制性内切酶位点的选择设定、筛选标记,都有精细的设计。因此,根据设计需要可以很方便的选择合适载体,进行 DNA 的体外重组。

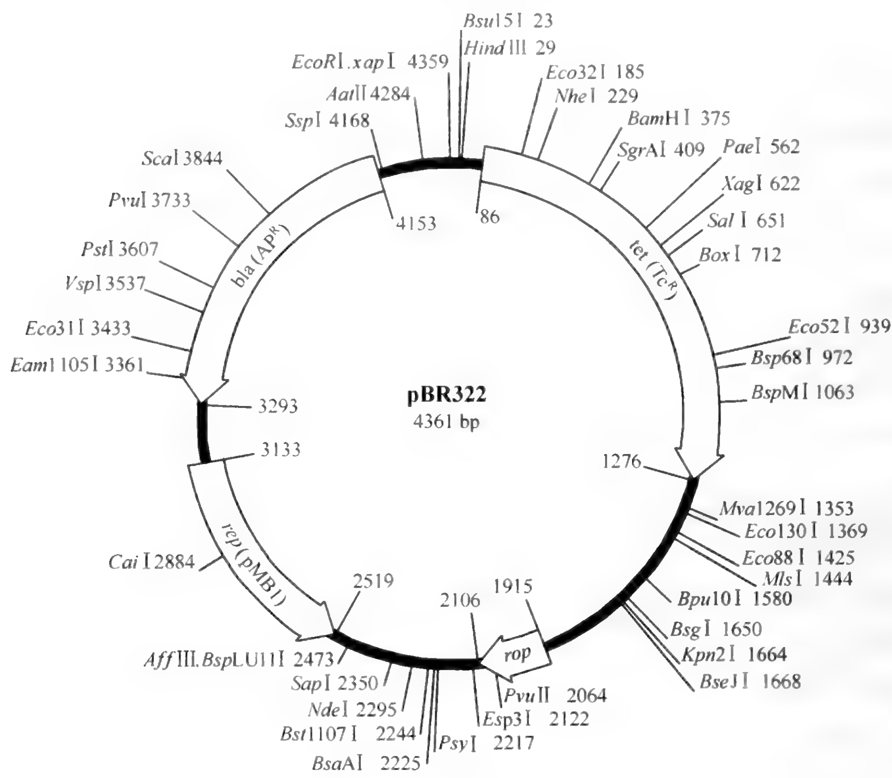


图 2-2 载体 pBR322 示意图

三、体外 DNA 重组

(一) 重组的概念

DNA 重组是指不同来源的 DNA 片段共价连接、通过重新组合,构成了具有两个 DNA 分子遗传信息的新重组体 DNA (recombinant DNA)。DNA 重组通常是指外来 DNA 片段(目的基因)与载体(质粒、噬菌体或病毒 DNA)的共价连接。

(二) 目的 DNA 与载体的连接

目的 DNA 及载体 DNA 经限制性内切酶水解后,可通过以下形式进行连接:

1. 黏末端连接 黏性末端(cohesive ends or stick ends)是指 ds DNA 分子经酶切割后所产生的限制性片段的单股末端。"黏性"指它能与另一个互补的单股末端配对结合为双链的性能。当以同一种Ⅱ型限制酶切割供体 DNA 及载体 DNA 时,可产生相同的单股黏性末端,如图 2-3 所示,以 *Eco*RⅠ酶切后末端都可为 TTAA。这样,供体与载体的 DNA 具有相同的互补的黏性末端,很易产生重组 DNA 分子。由于该重组 DNA 上保留了 *Eco*RⅠ的识别序列和切点部位,故有利于质粒扩增后目的基因的回收。另外,当单股黏末端超过 4 个碱基序列时,重组可仅靠互补的黏性末端碱基配对连接,可不需连接酶。

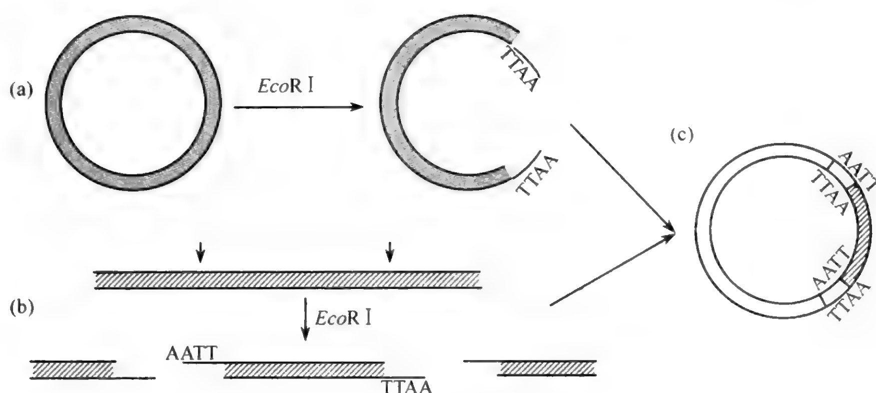


图 2-3 黏末端连接

两个 DNA 分子经同一酶切所形成的黏末端,因存在着互补的碱基序列可互相配对而结合,称为退火(anneal)。当一个 DNA 分子片段足够大时,自身也可退火形成环状(如载体 DNA 与载体 DNA 碱基配对而形成双连体重组分子)使目的 DNA 不易插入,形成无效重组分子。

黏末端连接主要有同源黏末端连接与定向克隆两种方法(详见第六章)。

2. 同聚物核苷酸末端法 该法是用末端核苷酸转移酶在 DNA 片段上人工制造一个黏末端,如在外来 DNA 片段 5′及 3′末端分别加上一小段单股同聚物尾(即一连串相同核苷酸),如 poly dA(或 poly dG),在载体两末端各加一段 poly dT(或 poly dC),这样两个 DNA 片段即可利用末端的 poly dA:poly dT(或 poly dG: poly dC)形成碱基配对,然后以连接酶结合(图 2-4)。若互补序列较长,可不需连接酶。该法的优点为不会发生自身环化。

3. 平齐末端连接法 T4 DNA 连接酶也可把只有平齐末端的两个 DNA 片段共价连成新的 DNA 分子,但其效应仅为黏性末端连接效率的百分之一。RNA 连接酶能提高 DNA 连接酶的连接效率,但其机制尚不清。

4. 人工接头连接 对平末端的 DNA 或没有互补黏末端的 DNA,除同聚物加尾法外,也可先连上人工设计合成的脱氧寡核苷酸双链接头,使 DNA 末端产生新的限制性酶切位点,经内切酶割后,即可按黏性末端相连。

5. T-A 克隆 T-A 克隆是用来直接克隆 PCR 产物的简便方法,也可连接 DNA 片段用于直接 DNA 测序(详见第六章)。

四、重组 DNA 的转化

(一) 转化

把以质粒为载体构建的重组 DNA ,在一定条件下引入受体细胞的过程称转化(transformation)。以噬菌体或病毒构建的重组 DNA 引入受体细胞的过程称转染(transfection)。接受了重组 DNA 的受体菌称为转化子或工程菌。

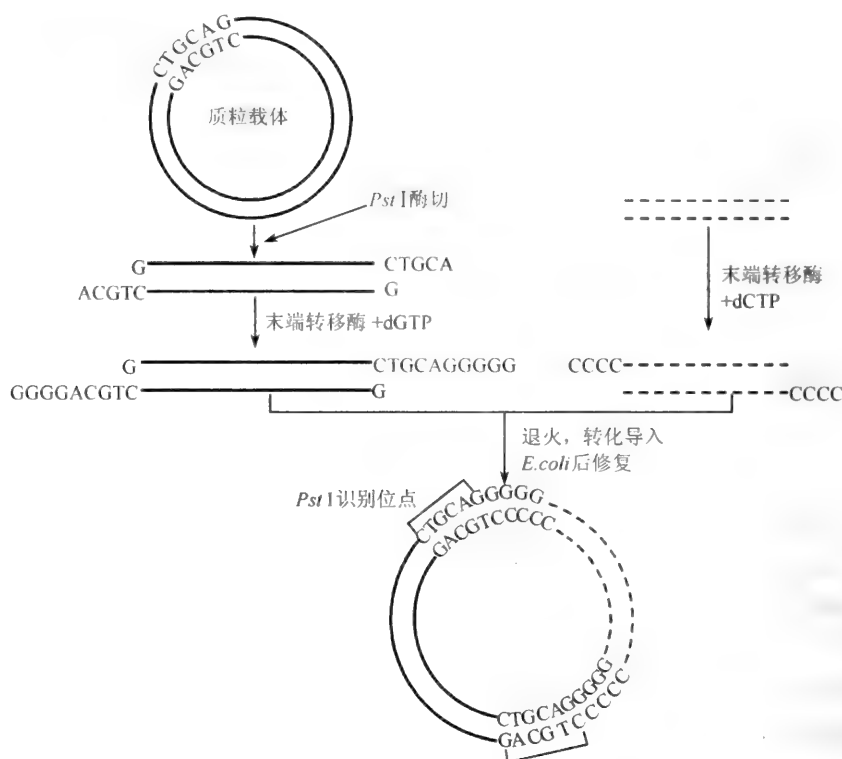


图 2-4 同聚物核苷酸末端连接法

(二) 感受态细胞及其特点

受体细胞处于感受态是转化成功与否的关键之一。感受态(competent)是指细胞处于最适于摄取和容忍外来 DNA 的生理状态。感受态细胞的特点为：

(1) 细胞表面暴露出一些可接受外来 DNA 的位点(以溶菌酶处理,可促使受体细胞的接受位点充分暴露)。

(2) 细胞膜通透性增加(Ca^{2+} 处理,可使膜通透性增加,使 DNA 直接穿过质膜进入细胞)。

(3) 受体细胞的修饰酶活性应表现最高,而限制酶活性最低,使转入的 DNA 分子不易被切除或降解。

(4) 受体细胞本身应处于非生长繁殖阶段(即受体细胞染色体相对稳定)。

基因工程中多采用限制酶阴性、修饰酶阳性($r^- m^+$)的大肠杆菌 C600、HB101、JM109 等作为受体细胞,比较容易接受重组 DNA 使其稳定扩增或表达。

(三) 转化方法

重组 DNA 转化的方法很多,列举常用的钙转化法及电转化法如下：

1. 钙转化法　大肠杆菌是应用最广的受体细胞之一。转化过程中受体细胞在 4℃ 以 Mg^{2+} 洗涤,然后加入 70～80mmol/L 的 Ca^{2+} 处理的受体细胞,随之加入重组 DNA,以 42℃ 瞬间热冲击处理,然后加入培养基,经培养后在培养板上筛选出相应的转化子。

氯化钙转化法由 Cohen 等人首创，其转化率一般为 $10^5 \sim 10^6$ 转化子/μg DNA。

2. **电转化法**　电转化法由 Dower 等人于 1988 年首次取得成功，目前 DNA 电转化仪已得到推广应用。

该法以高压脉冲电击细胞，使细胞摄取外源 DNA。电转化法不需制备感受态细菌，操作简便，适用于各种细菌、酵母菌以及真核细胞等的转化，其转化率可达 $10^9 \sim 10^{10}$ 转化子/μg DNA，转化率受电场强度、电脉冲时间长度及 DNA 的浓度影响。该法的缺点是转化细胞受强电场作用活性受到一定影响。由于电转化法转化率高，操作方便，故目前已广泛应用。

五、重组体的筛选鉴定与克隆扩增

(一) 重组体的筛选鉴定

感受态的受体细胞经体外重组 DNA 转化处理后，有一部分细胞获得了重组 DNA 而成为转化细胞(重组体)，但在整个反应体系中，尚存在着大量未转化的细胞，因此，进行克隆扩增前需设法将含有重组 DNA 分子的重组体从细胞群中分离出来，该过程即为重组体的筛选。

不同的克隆载体及相应的受体细胞，其重组体的筛选，鉴定方法不同。常用筛选重组体的方法有：平板筛选、电泳筛选、原位杂交筛选等，对重组体的进一步鉴定常用各种核酸分子杂交、免疫化学及核酸测序等方法。

重组 DNA 转化细胞后，首先进行平板筛选，确定转化效率，以利于进一步鉴定阳性重组体。常用方法介绍如下：

1. **平板筛选**　平板筛选是阳性重组体进行初步筛选的方法，简单而快速。筛选是依据载体 DNA 分子携带的筛选标记所赋予受体细胞在平板生长的表型特点进行的，供筛选用的载体 DNA 的遗传标志，在最初的设计时就应该考虑到。如质粒 pBR322 多采用抗生素抗性基因插入失活法来选择转化细胞。所谓插入失活(insertional inactivation)是在载体分子基因编码顺序中的限制性内切酶作用位点上，插入外源性 DNA 后，其编码供筛选的标记性产物或功能不能正常表达，此现象常作为一种筛选重组体的方法。

抗生素平板筛选为最常用的平板筛选法之一，因为多数克隆载体常有抗生素抗性基因，如抗氨苄西林基因(amp^r)，抗四环素基因(tet^r)等。载体 pBR322 具备 amp^r 及 tet^r 两个耐药基因，当选用 *Hin*d Ⅲ酶切时，外源性 DNA 正好插入到 tet^r 基因的位置，结果重组体表现为 tet^s，而 amp^r 不受影响。上述 pBR322 与外源 DNA 重组后转化 *E. coli* 时，利用插入失活、双抗生素对照筛选十分方便(图 2-5)。转化后可能产生以下三种细胞：

(1) amp^s、tet^s：重组 DNA 没有进入受体细胞，受体菌(*E. coli*)没有获得抗性基因，故不能在 Amp、Tet 平板生长。

(2) amp^r、tet^r：质粒 pBR322 进入了受体细胞，但质粒中没插入目的基因(即重组体没有进入受体细胞，进入的仅有载体质粒)，故保持了 amp^r、tet^r。

(3) amp^r、tet^s：重组质粒进入受体细胞，由于目的基因的插入使重组质粒 pBR322 的 tet^r 受到干扰而灭活，表现对四环素敏感，但因保持 amp^r 活性，仍可在 Amp 板上生长，故只有此组细胞是含有目的基因的转化子，应从平板上选择出来进行培养扩增。

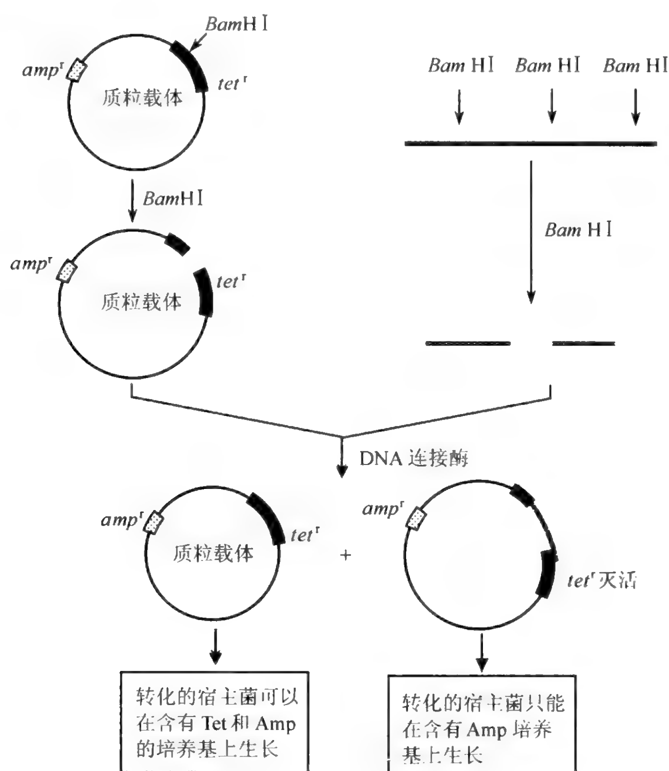


图 2-5 插入灭活示意图

平板筛选有多种方法，包括普通抗生素平板法，插入失活抗生素平板法、插入表达抗生素平板法及显色平板法等(详见第六、七章)。

2. DNA 限制性内切酶图谱分析与电泳筛选法 药物平板法筛选的重组体常有假阳性转化菌落，如自我连接载体，缺失连接载体等，仅用抗生素平板法不易鉴别，但采用电泳法可将这些假阳性菌落淘汰。电泳筛选法适用于插入片段与载体 DNA 片段相差较大的重组子的初步筛选，因为由这些转化菌落所抽提的重组质粒经酶切后产生的 DNA 分子大小不同，其 DNA 在凝胶中位于不同的区带，很易鉴别。但如果插入片段是与目的 DNA 或载体 DNA 大小相似的非目的 DNA 片段，则电泳法亦不能鉴别排除。这时可采用核酸分子杂交方法，即以目的基因片段制备的探针与重组 DNA 杂交，才能最终确定真正阳性重组体。

3. PCR 筛选重组体 有些载体的外源 DNA 插入位点两侧，存在恒定的碱基序列，因此，可设计与插入片段两侧 DNA 互补的引物，对小量抽提的重组质粒 DNA 进行 PCR 产物分析，此法不仅能迅速扩增插入片段，而且还可以进行 DNA 序列分析。如为 PCR 筛选测序而设计的 T-A 克隆载体已广泛应用(见第九章)。

以上方法仅是对重组体的初步筛选，作为重组体的鉴定则要求以更准确的方法直接证实或判断目的 DNA 的存在，如 DNA 测序、核酸分子杂交、免疫检测法(以单克隆抗体对表达产物的检测)、核酸电镜技术等。

4. 核酸分子杂交鉴定 核酸分子杂交(hybridization)是指两个不同来源的含有互补核

苷酸序列的单股核酸分子，通过碱基配对形成一个新的、稳定的双股分子的过程，是筛选、鉴定重组体最常选用的方法之一，按照参与杂交的核酸分子的不同，可分别形成 DNA-DNA、DNA-RNA、RNA-RNA 分子杂交。在核酸分子杂交鉴定重组体时必须具有相应的核酸探针进行检测。

所谓核酸探针是指以放射性同位素或非放射性标记物如地高辛、生物素等标记的一段特定已知序列的 DNA 或 RNA 片段，用于核酸杂交技术以检测出与其互补的待测酸分子。根据杂交的特点，分子杂交又可分为点杂交、原位杂交、Southern 杂交等(见第十章)。

菌落原位杂交技术是将带有重组子的菌落复印在硝酸纤维滤膜上，原位溶菌后，经 DNA 变性使其在原位固定于滤膜上，然后以相应的 DNA 探针通过分子杂交鉴定阳性重组体，最后对照参考平板上的相对位置，选择出相应的菌落。

5. DNA 测序　为获得有高效表达功能蛋白的重组体，检测插入片段的顺序及方向的正确性非常必要。经初步筛选后，一般需进行 DNA 测序。DNA 自动测序是目前已公认的确定重组体一级结构最可靠的技术。

(二) 重组体的克隆扩增

从转化细菌中筛选出含有阳性重组子的菌落经鉴定其正确性，即可通过细菌培养大量克隆扩增，以获得所需目的基因片段的大量拷贝或获得相应目的基因表达产物。如果实验设计的目的是制备纯化 DNA 片段(如制备探针所需 DNA 片段)，那么收获扩增后的细菌(工程菌)经质粒抽提、酶切、电泳即可回收目的 DNA。多数基因工程的重要目的在于实现目的基因的表达，以获得功能性蛋白产物，因此，目的 DNA 受体菌的表达是基因工程技术中最后也是十分关键的一步。

六、目的 DNA 在受体细胞中的表达

为使克隆的外源基因高效表达，必须有一个合适的表达系统以保证基因产物高效表达。而基因表达是在各种调控因子控制之下实现的，因此，必须精心设计构建高效表达载体系统以获得具有活性的蛋白质产物(见第七、十四章)。

目的 DNA 能否高效表达，除与表达载体自身特点及诱导条件有关，如温度、pH、诱导物的存在等，还取决于受体细胞容许目的 DNA 表达的多个层次：

1. 转录水平上启动子与受体细胞 RNA 多聚酶的统一问题。
2. 翻译水平上 mRNA 的核糖体结合部位与受体细胞核糖体的统一问题。
3. 外来 DNA 片段插入方向对表达的影响。
4. 其他，如转录后修饰、翻译后修饰等。以上任何一个过程均能直接影响到目的 DNA 的表达。

基因工程中目的 DNA 是以复制子作为载体，所以转化后通过复制子的自主复制才能扩增目的 DNA。但来自异种个体的 DNA 能否在受体细胞正确表达、转录、翻译出活性产物，这是关系到整个工程是否真正有价值的关键。

当以大肠杆菌为受体细胞时，它对各种来源的外来基因的容忍力均较强。枯草杆菌、葡

萄球菌质粒均可以在大肠杆菌中正常表达，但反之却不行。另外，外来基因在受体细胞中能否表达也与供体－受体的亲缘关系有一定的相关性。

对真核细胞表达载体的研究进展很快，特别是腺病毒表达载体已进入临床应用阶段。腺相关病毒载体、双启动子酵母表达载体、能够供真核与原核共表达的穿梭载体以及各种融合蛋白表达载体的使用，比较方便地实现了外源基因在真核细胞内的表达（见第七、八章）。

对目的 DNA 的表达产物尚需鉴定其产量和活性。表达产物的检测是确定整个设计是否达到预期目的的重要步骤，包括表达产物特异性及表达产物生物活性两方面的鉴定。特异性鉴定可采用免疫学方法，如免疫荧光素标记抗体、免疫沉淀、免疫印迹（Western blot）等；对蛋白质表达产物生物活性的鉴定则应根据其特点选择相应的检测指标或方法。

基因工程的流程设计参见图 2-6。

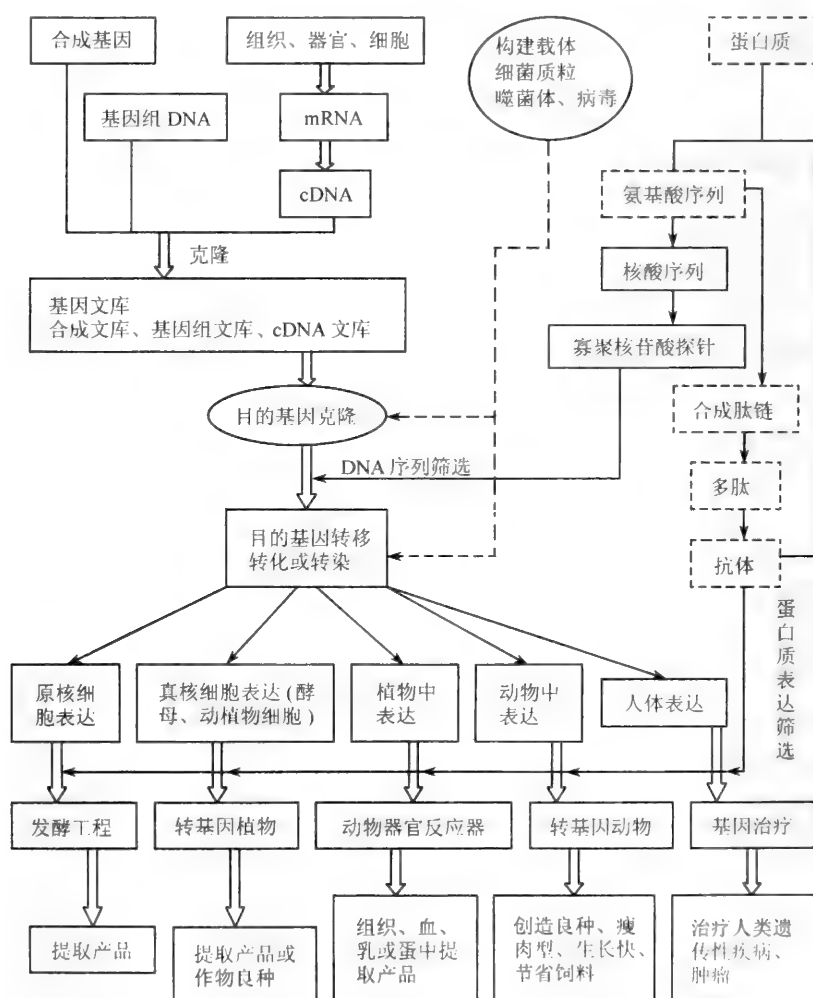


图 2-6　基因工程的流程设计图

（孙汶生）

第三章 基因工程的发展与应用

Chapter 3. Development and Application of Genetic Engineering

第一节 基因工程发展的历史背景

1953 年，Nature 上发表了伟大的科学家 Crick 和 Watson 提出的 DNA 双螺旋结构模型，揭开了长期困惑人们的生命之谜，宣告了分子生物学的诞生。它与爱因斯坦提出的相对论及海森堡创立的量子力学、量子场论共同被誉为 20 世纪最伟大的三项发现。

1990 年，由美国启动的人类基因组计划(human genome project, HGP)是探讨人类全部基因遗传信息的跨国科学工程，与制造原子弹的曼哈顿计划、阿波罗登月空间计划共同被誉为 20 世纪最伟大的三项科学计划。

基因工程学问世虽仅有几十年的历史，但很多先驱科学家及其开创性的科学研究为基因工程的发展建立了功不可没的丰功伟绩(表 3-1)。基因工程学为完成人类基因组计划提供了主要的技术并奠定了雄厚的理论基础，已成为生命科学的重要支柱学科之一。

基因工程学的崛起与发展基于该学科在理论与技术上的三大重大突破。

表 3-1 基因工程发展的百年回顾

年 代	人 物	事 件
1865 年	奥地利学者 G. Mendel	孟德尔遗传定律，遗传因子的概念
1869 年	F. Miescher	首次从莱茵河鲑鱼精子中分离到 DNA
1909 年	丹麦 W. Johannsen	首次提出以"gene"这一名词表达孟德尔的遗传因子概念
1910 年	美国遗传学家 Thomas Hunt Morgan	遗传学规律(基因连锁和互换定律)，首创基因学说
1940 年	Beadle 和 Tatum	蛋白质由基因编码
1944 年	3 位美国科学家	分离出细菌的 DNA，发现 DNA 是携带生命遗传物质的分子
1951 年	Pauling	蛋白质的 α 螺旋，蛋白质的二级结构
1953 年	J. D. Watson 和 F. H. C. Crick	在英国《自然》杂志发表"核酸的分子结构——脱氧核糖核酸的一个结构模型"，标志着 DNA 双螺旋结构的建立(1962 年获诺贝尔生理学或医学奖)
1957 年	A. Kornberg	从大肠杆菌中发现 DNA 聚合酶 I
1958 年	M. Meselson 和 F. W. Stahl	DNA 的半保留复制
1961 年	Crick	DNA、RNA、蛋白质的中心法则

续表

年　代	人　物	事　件
1966 年	M. W. Nirenberg 等	全部破译 64 个遗传密码，一部生物学的无字天书——密码字典问世，揭示了生物遗传信息的传递方式和规律
1970 年	Tem 和 Baltimore	病毒的反转录酶，修正了遗传信息流的中心法则
1972 年	P. Berg	划时代的第一次 DNA 重组(SV40 + λ 噬菌体)(1980 年获诺贝尔化学奖)
1973 年	Cohen 和 H. Boyer	划时代的第一次分子克隆，宣告了基因工程的诞生
1976 年	W. Gilbert 和 A. Maxam	化学直读法测 DNA 序列(1980 年获诺贝尔化学奖)
1977 年	Sanger	双脱氧链终止法测定 DNA 序列(1980 年获诺贝尔化学奖)
1977 年	Bergert	发现断裂基因(interruption gene)，揭示了真核基因的镶嵌结构，以及成熟 mRNA 加工机制——splicing
1977 年	Beyer	大肠杆菌第一个基因工程产品——人生长激素释放抑制素(somatostatin)
1978 年	Boyer 和 Itakura	大肠杆菌表达人生长激素基因
1981 年	R. D. Palmiter 和 R. L. Brinster	第一个转基因动物"超级老鼠"
1982 年	美国公司	第一个基因工程胰岛素商品化(Humulin)
1985 年	Kary B. Mullis	发明聚合酶链反应(PCR)
1985 年	Robert Sinsheimer	在加利福尼亚大学主持会议讨论人类基因组测序问题
1986 年	Muller	端粒(telomere)
1987 年	Dhundale	RNA 和 DNA 杂合子 msRNA，反转录单元和病毒的祖先 Retron，再次揭示生命的起源
1988 年	J. D. Watson	负责举世瞩目的人类基因组测序工作
1989 年	Greider	四膜虫的端粒酶(telomerase)
1990 年	Levine	明星基因——抑癌基因 *p*53
1990 年		第一个转基因玉米及转基因小麦植株诞生
1990 年 10 月		被誉为生命科学"阿波罗登月计划"的国际人类基因组计划启动
1994 年		提前一年完成人类基因组遗传图
1994 年		基因工程西红柿在美国上市
1995 年		英国《自然》杂志汇集发表了人基因组全物理图
1995 年	Chittenden, Farrow 和 Klefer	抑癌基因 *bak*
1996 年		*Methanococcus jannaschii* 基因组测序完成，确定地球上存在第三种生命类型
1996 年		DNA 芯片
1996 年	Morin	人 HeLa 细胞的端粒酶
1997 年	Jacobson	细胞凋亡(apoptosis)
1997 年	英国威尔姆斯(Weilmos)	克隆羊多利(Dolly)
1998 年 4 月		HGP 完成 50%
1998 年	美国 ONYX 公司	Bischoff 缺陷型腺病毒治疗癌症
1998 年		一批科学家在美国组建遗传公司，与 HGP 展开竞争
1998 年	夏家辉等	发现神经性耳聋基因

续表

年　代	人　物	事　件
1999 年		美国普及转基因玉米、大豆和棉花，种植面积 50%
1999 年 9 月		中国加入人类基因组计划，负责测定人类基因组全部序列的 1%
1999 年 12 月 1 日		人类第 22 号染色体测序完成
2000 年 4 月		中国科学家按照 HGP 的部署，完成了 1% 人类基因组的工作框架图
2000 年 5 月		人类第 21 号染色体测序完成
2000 年 6 月 26 日		科学家公布人类基因组工作草图，标志着人类在解读自身"生命之书"的道路上迈出了重要一步
2001 年 2 月 12 日		中、美、日、德、法、英等 6 国科学家和美国塞莱拉公司联合公布人类基因组图谱及初步分析结果
2001 年	J. M. Claverie	提出人类结构基因可能仅有 30 000 个的推论
2001 年 12 月		人类第 20 号染色体测序完成
2002 年 12 月 5 日		小鼠基因组序列草图完成
2003 年 1 月		人类第 14 号染色体测序完成
2003 年 4 月		人类基因组计划宣布完成

一、基因工程发展的关键性理论

1. 20 世纪 40 年代，Avery 和 Griffith 报道了肺炎链球菌的转化研究，首次证明了生物的遗传物质是 DNA。

2. 20 世纪 50 年代，Watson 和 Crick 提出了 DNA 结构的双螺旋模型，从此为搞清生物遗传物质的分子机制奠定了基础。

3. 20 世纪 60 年代，Nirenberg 等确定了遗传信息的传递方式，即 DNA→RNA→蛋白质遗传信息流的中心法则，全部破译了氨基酸的 64 个遗传密码，编排了一本震惊世界的密码字典，由此揭开了生物遗传的谜底，为基因工程学的诞生奠定了雄厚的理论基础。

二、基因工程的重大技术发明

基因工程三大技术的发明，为基因工程学的诞生建立了功不可没的丰功伟绩，提供了必不可少的关键性技术手段。

1. 限制性内切酶的分离与纯化　1970 年，Smith 和 Wilcox 首次从嗜血流感杆菌中分离纯化了限制性内切酶（restriction endonuclease, RE）*Hin*d Ⅲ，由此，为基因工程制备了精密的基因操作"手术刀"（见第四章）。

2. 反转录酶的发现　1970 年，Baltimore 和 Temin 同时各自发现了反转录酶，从而打破了传统的中心法则，并为真核基因的制备技术奠定了基础。

3. 载体技术　载体（vector）是指可容忍外源性 DNA 片段插入、可在细胞间转移并能在细胞内自主复制的 DNA 分子。细菌的质粒、噬菌体及某些病毒均有独立复制 DNA 的能力，并能通过一定方式进入细菌或受体细胞，故都是基因工程中良好的载体。Cohen 首次将

质粒作为基因工程的载体,为实现 DNA 体外重组提供了关键性的技术方法和有重要价值的实验材料,这是基因工程诞生的第三个技术准备。

1972 年,美国斯坦福大学的 Berg 首次实现 DNA 重组。其后,Cohen 于 1973 年进一步实现了重组 DNA 的转化,成为基因工程历史上第一个成功实现克隆转化的创举。因此,1973 年被确定为基因工程诞生的元年,Cohen 被誉为基因工程的创始人。20 世纪 70 年代,随着基因工程理论的突破与技术的发展,基因工程学水到渠成迅速发展为一门独立的工程技术学。

第二节 基因工程的发展与应用

基因工程的诞生,重组技术的兴起,开辟了按照人们的意愿短期内定向改造生物遗传性,创造新生物的新纪元。自 20 世纪 70 年代以来,基因工程已展现了惊人的发展实力和极其广阔的应用前景。

一、基因工程疫苗与 DNA 疫苗

(一) 基因工程疫苗

基因工程疫苗是指利用基因重组技术表达的重组蛋白作为疫苗,又称重组抗原疫苗(recombinant antigen vaccine)。制备基因工程疫苗的基本过程是将编码抗原免疫原性的基因与表达载体连接,导入细菌、酵母或真核细胞中,表达相应的蛋白质,接种给人体后可刺激机体产生有效的免疫应答。此类疫苗安全有效、成本低廉,已广泛应用于疾病预防。

(二) DNA 疫苗

DNA 疫苗,也称核酸疫苗(nucleic acid vaccine, NAV),是指直接把带有目的抗原基因(免疫原性基因,例如病原体的一个抗原编码基因)插入表达载体,通过注射或粒子轰击等途径将重组 DNA 导入动物细胞或人体,使之在活体细胞内持续表达出天然的抗原物质——目的蛋白质,这些目的蛋白质经正确的糖基化修饰等加工处理后,与 MHC 抗原形成复合物并被递呈到细胞表面——能诱导产生特异性细胞免疫(CMI)及体液免疫(HI)的物质。最近研究的较为深入的免疫刺激 DNA 序列(immunostimulatory DNA sequence, ISS)、重组病毒载体疫苗(RVVV)等均属于 DNA 疫苗(见第十九章)。

二、基 因 诊 断

基因诊断也叫 DNA 诊断、分子诊断,是通过从患者体内提取样本用基因检测方法来判断患者是否有基因异常或携带病原微生物。目前,基因诊断检测的疾病主要有三大类:感染性疾病的病原诊断、各种肿瘤的生物学特性的判断、遗传病的基因异常分析。一般情况下,对感染性疾病的诊断相对容易,制备相应的探针即能达到准确诊断;但对肿瘤及遗传病进行

基因诊断时应考虑到疾病的病理生理的复杂性(如基因的多功能性、多种功能调节与相互作用)及人类遗传的多样性(如表型所反映的基因型往往是通过极其多样的环境基础和其他重要方式诱导或起作用的),故目前大多只能以排除法对疾病进行诊断,例如对一种遗传性神经系统错乱症——亨廷顿病,若基因检测结果为阴性,则可排除此病。或通过对某些有遗传病危险倾向的人的预测、评估(如癌症、糖尿病、心脏病、高血压等),使其通过采取相应措施而预防或减少病情发生。

人类基因组计划为正常的人类基因组的研究提供了一个序列参考,在后基因组研究中,将不断发现新基因、致病基因及变异基因。有的是引起某些疾病或其他异常的原因,如编码某种酶、蛋白质、因子等的基因缺失或异常;有的虽然与这些疾病有关,但不是其发生的直接原因,如某些疾病与 HLA 相关联。当深入了解了人类基因的缺失或异常相关性时,通过检测有关疾病的发病基因,就可以诊断和预测疾病的发生。如 *p53* 基因与近一半肿瘤的发生有关。

基因诊断制剂和技术方法也将更加专一和快速、准确。未来 10 年中,基因诊断将从目前占现代疾病检测中的 0.5%扩大到占全部诊断检测的 8%～10%。

三、基因治疗

(一) 基因治疗发展的三个阶段

从 20 世纪 80 年代至 20 世纪末,基因治疗的发展经历了三个阶段:

1. 20 世纪 80 年代初是基因治疗的“禁锢时代”,从学术界到宗教、伦理、法律各界,对基因治疗存在很大争议。直到 1989 年,美国 FDA 才同意将载体导入作为“基因标记”的临床试验,科学家们在临床前研究方面进行了大量工作和舆论准备。

2. 1990 年之后,基因治疗进入临床试验,短短数年之内就有上百个临床方案,带来了医学生物学领域的一片狂热,使人们感到基因治疗即将成为临床治疗的一种成熟的方法,使一些尚未成熟的方案过早的进入了临床试验,给基因治疗的应用带来了一些不良影响。

3. 1995 年,美国 NIH 主持了对基因治疗临床试验的初步评估,使基因治疗从狂热转入理性化的正常轨道。1999 年,以电脉冲 DNA 导入为代表的新技术的问世,标志着基因导入系统的重大突破,为基因治疗的应用开创了新的契机。

(二) 基因治疗研究现状

将正常的治疗性基因导入人体,替换或矫正有害基因,即可达到基因治疗的目的。基因治疗最明显的应用是治疗遗传性基因疾病。1990 年,美国国立卫生研究所(NIH)首次对一名患有腺苷脱氨酶(ADA)缺乏症的 4 岁女孩进行了基因治疗,并获得了肯定的疗效,其 T 淋巴细胞半衰期从 30～35 天延长为 3～5 个月,由此标志着人类基因治疗正式开始。目前已有 300 多家基因治疗公司,已有 900 多种临床方案正在研究,基因治疗的疾病已涉及各种恶性肿瘤、心血管病、代谢病、感染性疾病、遗传性疾病和艾滋病等。

近年基因治疗发展平稳。截至 2004 年 1 月,全世界已有 918 个方案进入临床试验,其中恶性肿瘤治疗方案居首位,为 608 个;几乎大多数恶性肿瘤都有基因治疗方案,如神经、妇

科、消化道、肺、皮肤、头颈部以及造血系统恶性肿瘤等。其次为单基因遗传病，方案 90 个。我国现有基因治疗方案 3 个：复旦大学薛京伦实验室于 1991 年 12 月 2 日首次进入临床试验的血友病 B 基因治疗方案；反转录病毒载体导入 *IL*-2 基因治疗非小细胞肺癌的方案；腺病毒介导的 *p53* 基因转移结合放射线治疗头颈部肿瘤。

（三）基因治疗的相关技术

1. 目前基因治疗中临床试验所用的载体系统以反转录病毒居首位，其次是腺病毒、裸 DNA、脂质体、痘病毒、病毒疫苗、单纯疱疹病毒、腺相关病毒方案，以及 RNA 转移等。

2. 进行基因治疗用的基因按所占比例大小有细胞因子、抗原、抑癌基因、自杀基因、缺陷基因、耐药基因、受体以及其他基因。

3. 基因治疗必须解决三个关键问题　基因导入系统、基因表达的可控性及筛选更多更好的治疗基因。

（1）高效的、靶向性基因导入系统：基因治疗的关键是靶向性问题，只有将治疗基因输送并进入特定的靶细胞，才能在该细胞中得到高效表达。如果治疗基因不能导入大多数肿瘤细胞，至少要求尽可能不进入或较少进入正常细胞。迄今使用的病毒性载体中，如反转录病毒、腺病毒在体内肿瘤基因治疗时，除直接注射入瘤体外，若用全身给药，很难达到期望治疗的作用。对恶性肿瘤的体内形式的基因治疗，当务之急是建立靶向性强的载体系统。体外曾有过许多尝试，尤其是受体靶向的非病毒载体系统，如上海肿瘤研究所顾建仁院士创立的四元复合体导入系统。

（2）外源基因表达的可控性：最近电脉冲介导的 DNA 转化技术的改进使许多基因，如分泌性蛋白的基因可导入肌肉，维持相当时间的表达，其量有可能达到发挥药效的水平。

最理想的可控性是模拟人体内基因本身的调控形式，但其难度极大，对于导入基因的载体系统提出了严峻的挑战。长片段 DNA 进入细胞及其整合，涉及整合的位点，最终的理想是实现定点的“knock in”。

（3）筛选有治疗价值的基因：临床试验的治疗基因集中于少数基因，对极大部分多基因疾病（如恶性肿瘤、高血压、糖尿病等）的致病基因还有待阐明。这将有赖于人类基因组计划，尤其是功能基因组学的发展，包括目前已知和未知功能的基因的表达调控序列的确定，以及相互作用规律的阐明。

四、基因工程药物

以基因工程技术制备药物将成为今后医药发展的主流。基因工程药物主要有以下几个方面的研究与开发。

1. 基因重组药物　基因重组药物是以人工重组 DNA 表达产物作为治疗性药物，其表达/生产系统为各种细胞，如大肠杆菌、酵母细胞，或哺乳动物细胞。自 1982 年世界上第一个基因重组药物“人胰岛素”在美国上市以来，至今已有上百种产品问世，另有 300 多个品种处于临床试验阶段。全世界基因重组药物的使用以 16% 的速度增长。各国纷纷加大对该领域的投资，加强对其源头、生长点、制高点——功能基因的研究，并开始利用蛋白质工程技

术,对现有的重组药物进行分子改造(如构建融合蛋白)。随着后基因组研究的推进,有望开发出更多的基因重组药物。

广义的基因重组药物(生物技术药物)还包括重组疫苗和诊断或治疗用的单克隆抗体。

2. 植物基因工程药物　植物基因工程药物即利用转基因植物技术,在植物细胞中表达特定基因药物。将一些病毒或细菌的有用基因转入马铃薯等植物中,可大量、低成本地生产出疫苗,或通过直接食用该转基因植物而获得免疫力,如 HBV 转植物基因口服疫苗的研制。我国中药的一些有效成分的开发是一个很广阔的基因药物研究空间,可通过关键酶对其代谢途径进行遗传操作,有目的地获得大量有效成分,去除或减少有毒成分。

3. 动物基因工程药物　动物基因工程药物指利用转基因动物生产基因药物。用于分泌相应基因的蛋白质产物的动物器官(如乳腺)称为生物反应器。如用转基因绵羊生产蛋白质酶抑制剂 ATT,其产率达到每升奶 35g,可大量生产,用于治疗肺气肿,目前已进入Ⅲ期临床试验。

4. 基因组药物　基因组药物是沿着从基因序列→蛋白质→功能→药物的途径研制新药。基因组和后基因组学提供了庞大的信息资源和结果,为新药的开发提供了优化条件,前景广阔。

五、基因工程新技术的开发与应用

(一) SNP

单核苷酸多态性(single nucleotide polymorphisms, SNP)是近年研究和发现新基因的新思路、新方法。SNP 是人群中的某种单碱基差异,在人群或其他动物种群中的遗传变异中占有很高的比重。据估计,在人类基因组上平均每 1kb 就有一个 SNP 存在。

SNP 作为一种数量多、分布广的遗传标记,在疾病的研究过程中能够更好地帮助我们找到新的基因和阐明其机制。几乎每个基因中或者和其邻近的区域都有相关的 SNP 分布,并且能够和这些基因一起稳定地遗传下去。所以,通过比较正常人和病人的 SNP 的分布方式和频率,就可以确定哪些 SNP 与疾病相关联。这种研究方法在多基因控制的复杂疾病,如癌症、心血管疾病、糖尿病及一些较为少见的简单疾病中都是很有效的。SNP 研究还可以发现新的药物作用的靶基因,用于遗传性疾病的产前诊断以及在临床上根据病人个体的 SNP 差异进行个体化用药(详见第十六章)。

(二) 生物芯片技术

生物芯片包括 DNA 芯片、抗原芯片、抗体芯片、细胞芯片和组织芯片等。狭义上讲生物芯片是 DNA 芯片的代名词,其中的 DNA 微点阵(DNA 有序排列)包括直径为 200μm 或更小的数百至数千个点。目前研究和应用最多的生物芯片是 DNA 芯片,抗体等芯片仍在发展之中。生物芯片即缩小了的生化分析器,通过芯片上微加工获得的微米结构与生化处理相结合,将成千上万个与生命相关的信息集成在一块厘米见方的氧化硅、玻璃或塑料等材料上而制成。

20 世纪 90 年代初开始进行研制，芯片技术得到了迅速发展。发展的趋势是研制出高度集成化的集样品制备、基因扩增、核酸标记及检测为一体的便携式生物分析系统，即所谓的“微缩实验室芯片(Lab-on-chip)”，使现有的许多烦琐、费时、不连续、不精确和难以重复的生物分析过程自动化、连续化和微缩化，既可用于基础研究(如上文所述，进行方便快捷的基因分析、药物筛选等)，又具有巨大的应用价值(如疾病诊断等)和经济效益(见第十二章)。

(三) 基因敲除与转基因动物模型的建立与应用

基因敲除(gene knockout)是以基因工程的方法造成动物某一有功能的特定基因的缺失，以研究该基因的功能活性。而转基因动物(transgenic animal)是将某一生物特定的基因转入另一生物体，以获得该基因在受体动物中的稳定复制与表达，利于研究该基因的功能。如 HBV 转基因鼠使宿主特异性(人)的 HBV 在小鼠体内复制、表达，为 HBV 生物活性、致病机制及抗 HBV 药物筛选提供了不可缺少的实验动物材料。基因敲除与转基因技术详见第十八章。

第三节 人类基因组计划与后基因组时代

一、人类基因组计划

1990 年，在美国正式启动跨世纪的人类基因组计划(human genome project, HGP)，以期破译人类自身遗传秘密。2000 年 6 月 26 日，参加人类基因组计划国际协作组的各个国家同时宣布，人类基因组核苷酸序列的“工作框图”已完成。这是生命科学的里程碑，也昭示着人类已从基因组时代开始步入后基因组时代。

人类基因组是指传递人类遗传信息的 23 对染色体(22 对常染色体 + XX 或 XY 染色体)及其所携带的全部基因。基因是染色体上特定的 DNA 片段。人类基因组包括细胞核基因组与线粒体基因组，其结构框架如图 3-1 所示。

人类基因组约含 3 万个基因，由约 30 亿对碱基组成。人类基因组计划的目的就是找出 30 亿对碱基在染色体上准确的位置排序。这项工作分两步实施：①绘制人基因组的全部基因连锁图谱，也称染色体图或遗传图，该图早在 1992 年已完成；②基因的 DNA 序列(碱基对)测定，由此完成人类遗传密码的基本数据，已于 2000 年 6 月 26 日按期完成。

二、模式生物的基因组计划

在人类基因组计划的影响下，发展出模式生物的基因组计划。模式生物是分子遗传学家用于研究生命奥秘的基本工具。模式生物基因组计划将为后基因组时代人类基因功能的阐释提供很好的参照系统，加快和加深对人类基因功能的理解与运用。

1995 年，第一个细菌基因组——流感嗜血杆菌的全基因组序列发表，其后陆续完成了 11 种与人类疾病相关的病原菌和 5 种与工业、基础研究有关的细菌的基因组全序列分析。1996 年 4 月，真核单细胞酵母的基因全序列完成；1998 年 12 月，第一个多细胞真核生物的

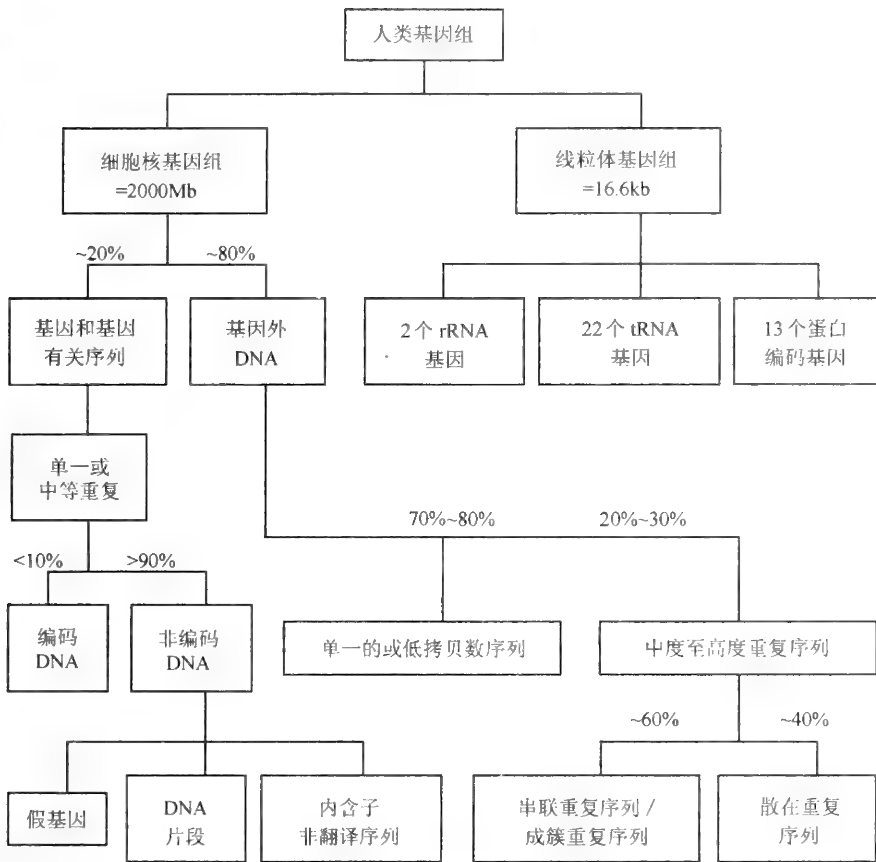


图 3-1 人类基因组框架示意图

基因组——线虫的基因组序列发表。果蝇基因组计划、水稻基因组计划正在进行之中，计划将于 2004 年完成模式植物拟南芥菜、于 2005 年完成模式动物小鼠的基因组计划。

三、人类后基因组计划

人类后基因组计划是由序列(结构)基因组学向功能基因组学的转移，即在基因组静态的碱基序列逐步搞清楚后，转而对基因组进行动态的生物学功能的研究。因为找出人类基因字母表的顺序仅仅是迈出了一小步，就 DNA 序列信息本身而言，并不能提供特定基因功能的确定信息。例如，一个给定的基因在什么确切位点、表达的时空调节及该期的确切功能如何？在这些海量的序列信息基础上，就可以在分子层面上探索人类健康和疾病的奥秘。基因是一种人类共有的有限的资源，基因产业能够创造巨大的商业价值，并将使人类生存的质量和境况发生革命性的飞跃。

随着后基因组计划的进行，近年来已衍生出许多新的分支学科，包括功能基因组学、结构基因组学、生物信息学、比较基因组学、蛋白质组学、整体生物学及药物基因组学等。

未来的基因组和后基因组学研究将为医疗保健带来新的模式，走进一个崭新的时代；将产生一个以特异的分子病理学为诊断依据，以患病前个人基因档案预测资料为依据的疾病预防式治疗模式。

（刘玉刚　孙汶生）

第二篇　基因工程学的基本理论与技术

Part Ⅱ. Basic Principles and Techniques of Genetic Engineering

第四章　基因工程工具酶及其应用

Chapter 4. Restricted Enzyme and Its Application

在基因工程中,无论是在特定部位准确切割 DNA 以获得所需的目的基因,还是将目的 DNA 连接于载体进行 DNA 分子重组,都离不开分子水平的精密"手术刀"、"缝纫针"——基因工程工具酶。基因工程中的工具酶是基因工程研究的技术工具,主要来源于细菌及真菌,种类繁多,功能各异。目前,分离并纯化的基因工程工具酶已超过 400 余种。其中较为重要的有十几种。本章将介绍几种常用的工具酶。

第一节　限制性核酸内切酶

宿主细胞中均存在着一对功能相反的酶——限制性核酸内切酶(restriction endonuclease, RE)和 DNA 甲基化酶或称修饰性核酸酶(modification nuclease),限制性核酸内切酶与其对应的甲基化酶对 DNA 底物有相同的识别顺序,但生物学功能是相反的。

限制性核酸内切酶简称限制酶,是一类能够识别 dsDNA 分子中的某特定核苷酸序列,并由此切割 DNA 双链结构的核酸内切酶,具有切除和降解无关外来 DNA 的作用,参与构成宿主抵抗外源 DNA 入侵的防御机制。限制性核酸酶内切 DNA 形成的片段,5′端为 P,3′端为 OH。

修饰性核酸酶简称修饰酶或甲基化酶,可对宿主本身的 DNA 进行限制性修饰,防止 DNA 被降解。

一、限制性核酸内切酶的命名原则

由于限制性核酸酶的数目非常庞大,需要有一个统一的命名法则。现介绍已得到认同的由 H. O. Smith 和 D. Nathans 于 1973 年提出的命名法。该规则以内切酶来源的微生物学名进行命名,具体法则如下:

1. 用产生酶的微生物"属"名的第一个字母(大写)及"种"名的前两个字母(小写),表示宿主菌的名称。如,埃希氏大肠杆菌(*Escherichia coli*)用 *Eco* 表示,嗜血性流感菌(*Haemophilus influenzae*)用 *Hin* 表示。

2. 若微生物有不同变种和品系,则需用变种或品系的第一个字母(大写)代表。如,大肠杆菌(*Escherichia coli*)R 株分离的酶用 *Eco*R 表示。

3. 同一微生物产生的几种限制酶,则根据其发现和分离的先后顺序用罗马数字表示。如大肠杆菌(*Escherichia coli*)R 株分离的第一种及第二种酶分别命名为 *Eco*RⅠ及 *Eco*RⅡ。

二、限制性核酸内切酶的分类

根据酶的识别序列及切割位点、催化条件及是否具有修饰酶活性,可将限制酶分为Ⅰ、Ⅱ、Ⅲ型。这三种不同类型的限制酶具有不同的特性,见表 4-1。

表 4-1 限制性核酸内切酶的类型及其特性

特点	Ⅰ型	Ⅱ型	Ⅲ型
限制和修饰活性	多功能酶	分开的核酸内切酶和甲基化酶	双功能酶
限制作用所需的辅助因子	ATP、Mg^{2+}、*S*-腺苷蛋氨酸	Mg^{2+}	ATP、Mg^{2+}、*S*-腺苷蛋氨酸
识别序列	*Eco*B:TGA(N)$_8$TGCT	旋转对称	*Eco*P1:AGACC
切割位点	在距识别序列至少 1000bp 处随机切割	位于识别位点或其附近	距识别序列 3′端 25bp 处
酶催转换	不能	能	能
甲基化作用的位点	识别序列	识别序列	识别序列
序列特异性切割	不是	是	是
在基因克隆中的作用	无用	非常有用	用处不大

注:N 代表任何一种核苷酸。

Ⅰ型限制酶:属复合功能酶,兼具限制、修饰两种功能,具有核酸内切酶、甲基化酶、ATP 酶和 DNA 解旋酶四种活性。Ⅰ型限制酶的活性发挥与 DNA 的甲基化程度有关,若识别位点上两条 DNA 链均未甲基化,就行使内切酶功能,并在切割 DNA 的同时或以后转变为 ATP 酶;若位点上只有一条链甲基化,则发挥修饰作用,使另一条链也甲基化;若位点上两条链均已甲基化,则与位点解离。Ⅰ型限制酶的显著特点是在 DNA 链上的识别序列与切割位点不同,一般在离识别位点 1kb 至几个 kb 的部位随机切割 DNA,需要 ATP 的水解作用提供能量。Ⅰ型酶无固定切点,不产生特异性 DNA 片段。如 *Eco*B、*Eco*K。

Ⅱ型限制酶:即通常所说的限制酶。具有这类酶的微生物其限制、修饰系统分别由两种不同的酶完成,即限制性内切酶和修饰性甲基化酶。Ⅱ型限制酶的特点是相对分子质量小,仅需 Mg^{2+} 作为催化反应的辅助因子,能识别并切割 DNA 链上的特异性核苷酸顺序,产生特异性 DNA 片段,因此,在基因工程中Ⅱ型限制酶的应用非常广泛。

常用的Ⅱ型限制酶及其识别序列与切割位点,见表 4-2。

Ⅲ型限制酶:与Ⅰ型限制酶相似,但其不同点是Ⅲ型限制酶的切割位点在识别序列的一侧约 25bp 处的核苷酸位置上单链切割 DNA 分子,有特异性的切割位点。此类酶数量相当少,如 *Eco*PⅠ。

Ⅰ、Ⅲ型限制酶对基因工程的意义不大,本节重点介绍Ⅱ型限制酶。

表 4-2　部分Ⅱ型限制酶的性质及酶切位点

酶名称	识别序列	缓冲液	温度/℃	异源同工酶	可互相连接的酶
Ava Ⅰ	C↓PyCGPuG	M	37		*Sal* Ⅰ或 *Sml* Ⅰ
*Bam*HⅠ	G↓GATCC	K	30		*Bcl* Ⅰ, *Bgl* Ⅱ, *Mbo* Ⅰ, *Xho* Ⅱ
Bgl Ⅱ	A↓GATCT	L	37		*Xho* Ⅱ, *Bam*HⅠ
Bcl Ⅰ	T↓GATCA	M	60		*Bam*HⅠ, *Xho* Ⅱ
*Bst*EⅡ	G↓GTNACC	M	60		
*Bst*ZⅠ	C↓GGCCG	H	50		
Dpn Ⅰ	G↓ATC	M	37	*Sau*3AⅠ	Blunt
*Eco*RⅠ	G↓AATTC	M/H	37		
*Eco*RⅤ	GAT↓ATC	H	37		
*Hin*dⅢ	A↓AGCTT	M	37～55		
*Hin*cⅡ	GTPy↓PuAC	M	37		任何平头末端酶
*Hin*fⅠ	G↓ANTC	M	37		
Hpa Ⅰ	GTT↓AAC	L	37		任何平头末端酶
Hpa Ⅱ	C↓CGG	L	37		*Taq* Ⅰ, *Cla* Ⅰ
Kpn Ⅰ	GGTAC↓C	L	37		
Mbo Ⅰ	↓GATC	M	37	*Sau*3AⅠ	*Xho* Ⅰ, *Bam*HⅠ
Nco Ⅰ	C↓CATGG	H	37		
Pst Ⅰ	CTCGA↓G	M	21～37		
Pvu Ⅱ	CAG↓CTG	M	37		任何平头末端酶
Sac Ⅰ	GAGCT↓C	L	37	*Sst* Ⅰ	
Sal Ⅰ	G↓TCGAC	M/H	37		*Ava* Ⅰ, *Xho* Ⅰ
*Sau*3AⅠ	↓GATC	M	37		*Bam*HⅠ, *Bgl* Ⅱ
Sma Ⅰ	CCC↓*GGG	K4	37	*Xma* Ⅰ	任何平头末端酶
Sst Ⅰ	GAGCT↓C	L	37	*Sac* Ⅰ	
Taq Ⅰ	T↓CGA	L	65		*Hpa* Ⅱ, *Asu* Ⅱ, *Cla* Ⅰ, *Msp* Ⅰ
Xba Ⅰ	T↓CTAGA	H	37		
Xho Ⅰ	C↓TCGAG	H	37		*Ava* Ⅰ, *Sal* Ⅰ
Xma Ⅰ	CCC↓GGG	L	37	*Sma* Ⅰ	某些 *Ava* Ⅰ
Xma Ⅱ	CTCGA↓G	M	37	*Pst* Ⅰ	

注：识别序列用一条链按5′→3′方向表述，箭头所指系酶切位点；Py代表嘧啶碱基C或T，Pu代表嘌呤碱基A或G；星号(*)代表该碱基被特定的甲基化酶所修饰；L、M、H分别代表该酶的缓冲液为低盐、中盐、高盐缓冲液。

三、Ⅱ型限制酶的识别与切割序列

（一）Ⅱ限制酶的识别与切割特点

1. Ⅱ型限制酶具有特定的识别顺序　绝大多数的Ⅱ型限制酶，都能够识别由4～8个核苷酸组成的特定核苷酸序列，即识别序列，Ⅱ型限制酶在其识别序列内切割DNA分子，

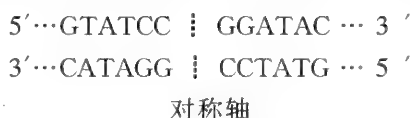
因此,识别序列又称为Ⅱ型限制酶的切割位点或靶序列。

Ⅱ型限制酶识别靶序列的大小决定着产生特定 DNA 片段的大小。如果 DNA 的碱基组成是均一的,限制酶位点在 DNA 链上的分布是随机的,则要求识别 4 个核苷酸靶序列的酶(如 *Hpa* Ⅱ、*Mbo* Ⅰ 等),在一个庞大的 DNA 链上,平均每 4^4 个核苷酸即可遇到一个靶序列($1/4^4 = 1/256$);而对要求识别 6 个核苷酸靶序列的酶(如 *Bam*HⅠ、*Eco*RⅠ、*Pst* Ⅰ 等),则平均每 6^6 个核苷酸才能遇到一个特异的识别序列(即 $1/6^6 = 1/4096$)。

2. Ⅱ型限制酶的识别序列具有 180 度的旋转对称性　Ⅱ型限制酶的识别序列呈完全回文结构。所谓回文结构指的是形似回文的一段 DNA 序列,其特点有二:

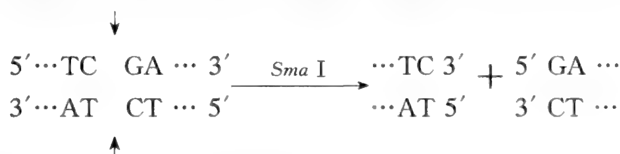
(1) 以某一对核苷酸为中轴,左右同数目的核苷酸彼此呈碱基互补;

(2) 这两股 DNA 链若按同方向阅读(如 5′→3′或 3′→5′),其核苷酸顺序相同。如:



3. Ⅱ型限制酶对 dsDNA 切割时,可产生两种不同的切口(末端)

(1) 平齐末端(flush 或 blunt end):Ⅱ型限制酶在其识别序列的对称轴上对 dsDNA 的两条链同时切割,两条链的断裂位置处在对称轴上,形成双股平齐末端。如 *Sma* Ⅰ 等。

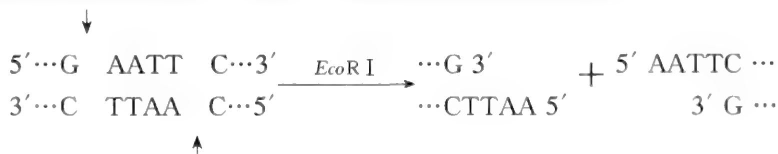


具平齐末端的 DNA 片段不易于重新环化。

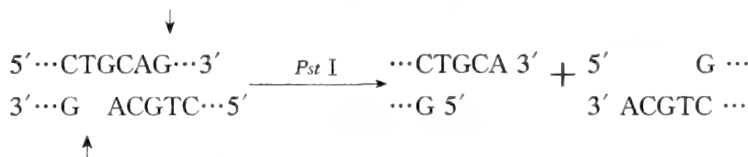
(2) 黏末端(cohesive end):大部分的Ⅱ型限制酶(如 *Bam*HⅠ、*Eco*RⅠ等)的切口作用点具 180°旋转对称性,形成黏性末端。该末端在双股 DNA 断口处,各产生一小股互补的核苷酸单股区,若此单股区有四个以上核苷酸序列,就可产生稳定的碱基配对。两个不同的 DNA 分子,经同一个限制酶切割所形成的黏末端是相同的,经互补的碱基配对很容易退火而组合成新的 DNA 分子。重组的 DNA 分子仍保留了原来的酶切序列,这对重组 DNA 分子中目的 DNA 的回收十分重要。由于黏末端具有 180°旋转对称性,故经酶切的 DNA 分子,也可自身重新环化。

黏性末端又有 5′-黏末端与 3′-黏末端之分:

1) 5′-黏性末端:该类酶在其识别序列的对称轴两侧,从 5′-P 末端切割 dsDNA 的两条链,产生 5′-P 单链延伸的黏性末端。如 *Eco*RⅠ 酶切后形成的黏末端:

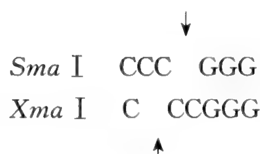


2) 3′-黏性末端:该类酶在其识别序列的对称轴两侧,从 3′-OH 末端切割 dsDNA 的两条链,产生具有 3′-OH 单链延伸的黏性末端。如 *Pst* Ⅰ 酶切后形成的黏末端:



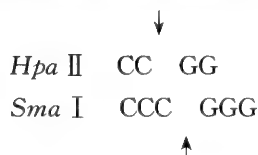
（二）有特殊性质或相互间有一定关系的Ⅱ型限制酶

1. 异源同工酶(isoschizomers)　也称为同裂酶，指来源不同识别序列相同的酶。该类酶切割DNA的位点或方式可相同，亦可不同。如 *Sma* Ⅰ与 *Xma* Ⅰ，这两种酶识别序列相同，但酶切位点不同：



又如，*Hpa* Ⅱ与 *Msp* Ⅰ，它们的识别与切割序列均相同(C↓CGG)，但 *Msp* Ⅰ还可以识别切割已甲基化的序列，如GG↓mCC。

2. 同尾酶(isocaudarner)　识别序列与切割位点相互有关的一类酶。它们来源各异，识别序列也各不相同，但都可以产生出相同的黏性末端。如：



Sma Ⅰ所识别的6个核苷酸序列中，包含有 *Hpa* Ⅱ识别的4个核苷酸序列，所以，*Hpa* Ⅱ能识别并切割 *Sma* Ⅰ的核苷酸序列，但反之则不行。又如：



上述四个酶的识别序列中都有GATC，酶切后均产生相同的黏性末端——GATC。但仅有 *Mbo* Ⅰ能识别并切割其他三个酶的识别序列，而其他三个酶之间却不能互相识别其核苷酸序列。

同尾酶产生的DNA片段，由于具有相同的黏性末端，因而，能够通过黏性末端之间的碱基互补而彼此连接，形成的位点称之为“杂种位点”(hybrid site)。但此类杂种位点不能够再被原来的任何一种同尾酶识别切割。

异源同工酶与同尾酶在基因工程中有一定的作用。由于消化条件和来源的限制，不能用一种酶消化某类底物时，则可用以上两种酶代替。

3. 远距离裂解酶 (distant cleavage)　此类酶在DNA链上的识别序列与切割位点是不一致的，它们在某一核苷酸区域与识别序列结合，然后滑行到识别序列以外的另一个位点进行切割，这一点与Ⅰ型酶相似，但不同的是其切点与识别位点的距离是一定的，而且不像Ⅰ

型酶那么遥远，一般为 10 个碱基左右。如：

Mbo Ⅱ GAAGANNNNNNNN↓（N 代表任何一种碱基），切点与识别位点相隔 8 个碱基；

Hga Ⅰ GACGCNNNNN↓，切点与识别位点相隔 5 个碱基。

此类酶在基因工程中具有一定的应用价值。

4. 可变酶 此类酶是Ⅱ型酶中的特例，其识别序列中的一个或几个核苷酸是可变的，该识别序列往往超过 6 个核苷酸。如：

*Bst*pⅠ：GGTNACC，识别 7 个核苷酸序列，其中 1 个碱基可变。

Bgl Ⅰ：$GCC(N)_4NGGC$，识别 11 个核苷酸序列，其中 5 个碱基可变。

四、影响Ⅱ型限制酶作用的因素

除 DNA 本身的因素外，反应混合液中的离子强度、酸碱度、试剂（甘油、DMSO）及酶浓度等因素，对酶的专一性和反应速度均有影响。

（一）DNA 对酶切反应的影响

1. DNA 的纯度 限制性核酸内切酶消化 DNA 的反应效率，在很大程度上取决于所使用 DNA 本身的纯度。若 DNA 纯化过程中污染了某些制剂，如酚、氯仿、乙醇、TE 以及高浓度的盐离子等，都有可能抑制核酸内切酶的活性。应用碱变性法小量碱法抽提的 DNA，常常含有这类杂质，一般采用以下三种方法，可提高核酸内切酶对低纯度 DNA 的消化效率。

（1）增加核酸内切酶的用量，平均每微克底物 DNA 可加入 5～10 单位酶。

（2）扩大酶切反应的体积，使潜在的抑制因素被相应的稀释。

（3）延长酶切反应的时间，可延长为 3～4h。

2. DNA 的甲基化程度 识别序列中特定核苷酸的甲基化修饰，会强烈地影响限制性核酸内切酶的活性，导致局部消化，甚至完全不被消化。在基因克隆中，使用失去了甲基化酶的大肠杆菌菌株制备质粒 DNA，可避免产生此类问题。

（二）反应缓冲液对酶切效率的影响

1. 离子强度、pH 的影响 限制性内切酶在非标准反应条件下，能够切割一些与其特异识别序列类似的序列，这种现象称“星号活性”（star activity）。星号活性一般在相应限制酶的名称右上角加一个星号（＊）表示。例如，当缓冲液中的离子强度保持低水平时，如 $MgCl_2$ 浓度由 5mmol/L 降至 2mmol/L（pH8.5）时，*Eco*RⅠ对底物专一性识别能力降低，其识别序列可从正常的 6 核苷酸（GAATTC）减到 4 核苷酸（NAATTN），识别的特异性发生“松动”，可从其“正确”识别序列以外的其他位点切割 DNA 分子。*Eco*RⅠ的这种星号活性，以 *Eco*RⅠ* 表示。

有几种二价金属离子能使限制性 *Eco*RⅠ变为 *Eco*RⅠ*，从高到低依次为：Mn^{2+}、Mg^{2+}、Zn^{2+}、Co^{2+}。

缓冲液为高 pH、低离子强度时较易诱发酶的星号活力出现。

2. 某些试剂可改变酶的专一性　实验证明，NaCl 能稳定 *Bsu* Ⅰ对其识别序列 GGCC 的专一性。但当降低 NaCl 总浓度和(或)增加 pH 时，也能促进 *Eco*RⅠ* 的活性。

另外，当疏水试剂如甘油含量过高(>10%，V/V)或二甲基亚砜等存在时，使 *Bam*HⅠ的识别与切割专一性发生改变，导致不能识别自己的正确序列。

(三) 酶对消化效率的影响

内切酶的量对酶切反应的影响较大。一般是 1μg 底物(λDNA)，在 50μl 终反应体系中，在推荐的反应缓冲液和温度下反应 1h，达到完全酶解所用的酶量为 1U。若酶解 1μg 底物，按此标准，1U 的内切酶可达到完全消化。酶解 1μg 以上的 DNA，可按此标准体系的比例进行放大。但在实际应用中，一般要用稍过量的酶(1μg 加 2～5 单位酶)反应较长的时间，才能够达到完全酶解。但是，不能过分加大酶的用量。

酶量过大对酶切反应的影响包括两个方面：其一，酶过量(一般为>100U/μg DNA)会影响酶识别序列的正确性，如 *Bam*HⅠ的正常识别序列为 G↓GATCC，当酶过量时，其识别序列变为 NGATCN，称为 *Bam*HⅠ*；其二，酶制剂含甘油成分，酶过量，会因为甘油量大而影响其识别的专一性，诱发星号活力。

(四) 酶切反应的温度

消化反应的温度，是影响限制酶活性的另一个重要因素。不同的限制性核酸内切酶，具有不同的最适反应温度。大多数限制性核酸内切酶的标准反应温度都是 37℃，但个别酶要求 37℃ 以外的反应温度(表 4-2)。酶切反应的温度高于或低于最适温度，都会影响核酸内切酶的活性，甚至最终导致酶完全失活。

(五) 酶解时间

酶解时间可以通过加大酶量而缩短，反之，酶量较少时可以通过延长酶解时间以达到完全酶解的目的。不同的 DNA 底物在一定酶量和一定时间内，其酶解效率不同，可以根据 DNA 底物上酶切位点的多少，与 λDNA 存在位点的数目进行比较后，再决定酶量和酶解时间。

五、限制性核酸内切酶在基因工程中的应用

DNA 分子的完全与正确酶解，对于 DNA 连接、基因克隆分子筛选和鉴定、DNA 序列测定以及构建基因组文库意义重大。

1. 局部酶切消化作用　用于 DNA 物理图谱的绘制。

理论上，限制性核酸酶对 DNA 分子的切割反应，应为完全酶切消化作用(complete digestion)。但实际上，由于组成 DNA 分子的 4 种核苷酸的含量不是等量的，其排列顺序也不是随机的，因此，限制性核酸酶对 DNA 分子的实际消化作用要低于完全消化的频率，此类不完全的限制酶消化反应，称之为局部酶切消化(partial digestion)。如果通过缩短酶切反应的时间，或降低反应的温度以约束酶的活性，就可以达到局部消化的目的，可用于绘制 DNA

的物理图谱。

2. 对基因组 DNA 的切割——构建基因组文库　基因组文库(genomic library)是含有某种生物体全部基因的随机片段的重组 DNA 克隆群体。构建基因组文库时,由于提纯的原核或真核细胞染色体 DNA 相对分子质量均较大(在 100kb 以上),需要限制性内切酶使之消化为一定大小的片段,再与相应载体连接,构建文库(详见第六章)。

3. 重组子的筛选鉴定　对于抗生素平板初步筛选具有重组子的菌落,小量培养后,分离出重组质粒 DNA,用相应的内切酶消化重组子释放出插入片段,然后凝胶电泳检测插入片段和载体的大小。对于可能存在双向插入的重组子,还可以用内切酶消化鉴定插入方向(详见第六章)。

此外,利用限制酶对 DNA 片段的完全消化作用,可用于 DNA 序列测定。两个不同的 DNA 分子,经同一个限制酶切割后,可用于 DNA 重组克隆及亚克隆(详见第六章)。

第二节　DNA 连接酶

DNA 重组技术的核心步骤是 DNA 片段的体外连接。DNA 连接的本质是一个由连接酶催化的生物化学过程。在这个过程中,连接酶(ligase)的活性对连接产物的形成甚为关键,因此,DNA 连接酶是体外重组 DNA 分子必不可少的基本工具酶之一。限制性核酸内切酶将 DNA 分子切割成不同大小的片段,而连接酶将不同来源的 DNA 片段组成新的杂种 DNA 分子,并把他们彼此连接并封闭起来。

一、DNA 连接酶的分类及特性

(一) T4 DNA 连接酶

T4 DNA 连接酶是由大肠杆菌(*E. coli*)T4 噬菌体 DNA30 编码的产物,最早是从 T4 噬菌体感染的大肠杆菌中分离纯化的,相对分子质量为 68×10^3,是一种单链多肽酶,能催化两个 DNA 片段的 3′-OH 末端与 5′-P 末端形成磷酸二酯键,需要 ATP 作为辅助因子并由 ATP 提供能量。T4 噬菌体 DNA30 现已被克隆,并可在大肠杆菌中大量表达,因此 T4 DNA 连接酶比较容易制备。T4 DNA 连接酶可高效催化 dsDNA 片段间黏末端的连接,并可催化平末端 DNA 片段的连接,因此,在分子生物学研究和基因克隆中应用非常广泛。

(二) 大肠杆菌 DNA 连接酶

又称为 *E. coli* DNA 连接酶,是由大肠杆菌染色体基因组中 *lig* 基因编码的,相对分子质量为 75×10^3。*lig* 基因已经被克隆,并能在相应大肠杆菌细胞中大量表达。此连接酶不能催化平末端 DNA 分子之间的连接,其底物只能是带缺口的双链 DNA 分子和具有同源互补黏末端的不同 DNA 分子,并以 NAD^+ 作为辅助因子。

二、连接酶催化DNA连接的机制

DNA连接酶可催化两相邻的DNA链中的一条链3′-OH端与另一条链5′-P末端形成磷酸二酯键，这个过程是一种耗能反应。大肠杆菌DNA连接酶是在大肠杆菌及其他细菌中催化DNA分子的连接反应，利用NAD^+（烟酰胺腺嘌呤二核苷酸，nicotinamide adenine dinucleotide）作为能源；而T4 DNA连接酶则在动物细胞及噬菌体中催化DNA分子的连接反应，利用ATP（腺苷三磷酸，adenosine triphosphate）作能源。二者除了能源辅助因子不同之外，其他的作用机制并无差别。

下面以T4 DNA连接酶为例，简单讨论连接酶催化连接的分子机制。

T4 DNA连接酶催化DNA连接的反应分为三步（图4-1）：ATP与T4 DNA连接酶形成复合物，通过ATP中的磷酸与T4 DNA连接酶中的亮氨酸的氨基，形成磷酸-氨基键而连接，即形成酶-ATP复合物（T4 DNA连接酶-亮氨酸-*N*-核糖-腺嘌呤）；酶-ATP复合物活化DNA链5′端的磷酸基团，形成磷酸-磷酸键；DNA链3′端的羟基活化并取代ATP，与5′端磷酸根形成磷酸二酯键，并释放出AMP，完成DNA链之间的连接。

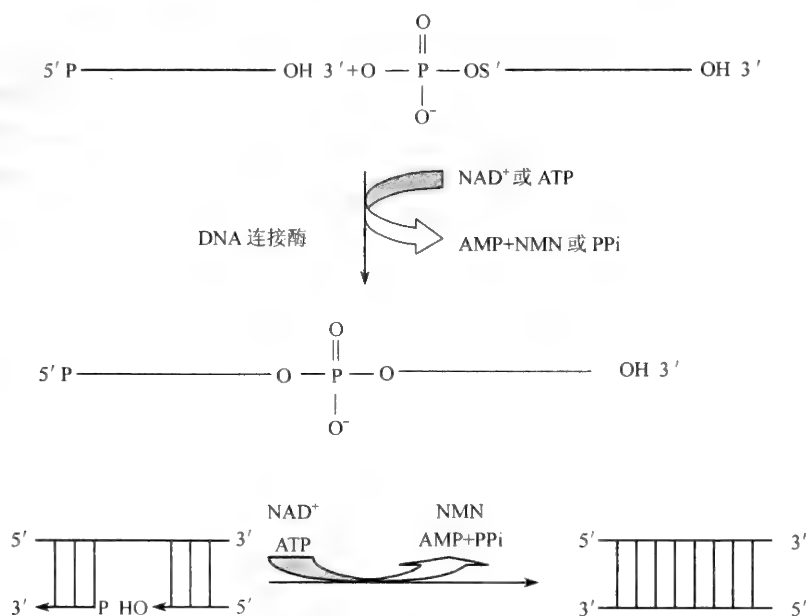


图4-1　T4 DNA连接酶催化DNA连接的分子机制

三、影响连接效率的因素

1986年，V. King和W. Blakeskey提出，以连接反应产物转化感受态细胞的能力，作为判断连接效率的标准。连接反应中的ATP浓度、连接酶浓度、连接反应时间、连接反应温

度、插入片段与载体分子的摩尔比值等参数，均可影响连接反应产物的转化效率。

1. 连接反应的温度　是影响转化效率的最重要参数之一。连接酶连接缺口DNA的最佳反应温度是37℃，但在此温度下，黏性末端之间的氢键结合不稳定。因此，连接黏性末端的最佳温度，应该保证黏性末端退火及连接酶活性、反应速率之间的平衡，一般以4～15℃为最佳。黏性末端中G+C含量多时，其连接反应温度可以增高。

2. 连接酶的浓度　在平末端DNA的连接反应中，最适的反应酶量大约为1～2U；而对于黏性末端(如*Eco*RⅠ)间的连接，在同样条件下，酶浓度仅为0.1U，便能达到最佳转化效率。在不同生产厂家，由于其所规定的连接酶的活性定义单位不同，因此，T4 DNA连接酶活性可相差几十倍，甚至上百倍。在实际应用中，应参照说明书确定使用酶量。

3. ATP浓度　反应浓度变动范围一般应保持在10μmol～1mmol之间。

4. 连接反应时间　当连接反应温度为12～30℃之间时，连接反应时间约为1～16h。若连接温度更低(如4℃)，连接时间应适当延长至24～48h。

5. 插入片段与载体分子的摩尔比值　插入片段与载体分子以2～3∶1的比例最为合适。

四、DNA连接酶在基因工程中的应用

1. 对同源黏末端DNA分子的连接　同源黏末端包括相同一种内切酶产生的黏性末端和不同的内切酶产生的互补黏性末端，后者连接成的DNA顺序不能再被原切割内切酶识别切割。此种连接反应的产物中，存在一定数量的载体自身环化分子，使转化菌中出现较高的假阳性克隆。在连接前将线性DNA分子进行5′端脱磷酸处理，可避免假阳性克隆的出现。

2. 对平末端DNA分子的连接　大肠杆菌DNA连接酶不能催化平末端DNA片段的连接，平末端DNA分子之间的连接，需要T4 DNA连接酶催化。此种连接方式，既可以恢复原始的酶切位点，又可以创建一个新的酶切位点，使某些重组子的筛选鉴定大大简化。平末端DNA分子连接的效率较黏末端低，加入适量的PEG能促进平末端的连接。

第三节　反转录酶

反转录酶即依赖于RNA的DNA聚合酶，是Temin等于1970年在致癌RNA病毒中发现的一种特殊的DNA聚合酶，该酶以RNA为模板，根据碱基配对原则，按照RNA的核苷酸顺序(其中U与A配对)合成DNA。这一过程与一般遗传信息流转录的方向相反，故称为反转录，催化此过程的DNA聚合酶叫做反转录酶(reverse transcriptase)，也称之为逆转录酶。后来发现反转录酶不仅普遍存在于RNA病毒中，哺乳动物的胚胎细胞和正在分裂的淋巴细胞中也有反转录酶。目前，已经从多种RNA肿瘤病毒中分离到这种酶，但最普遍使用的是来源于鸟类骨髓母细胞瘤病毒(avian myeloblastosis virus, AMV)的反转录酶及来源于莫洛尼鼠白血病病毒(Moloney murine leukemia virus, M-MLV)的反转录酶。

一、AMV 的生化特性

AMV 由 α 和 β 两条多肽链组成，其相对分子质量分别为 65×10^3 和 95×10^3。α 链具有反转录酶活性和 RNaseH 活性，RNaseH 活性是一种核糖核酸外切酶活性，能以 5′→3′或 3′→5′方向特异地降解 RNA-DNA 杂交分子中的 RNA 链，继而作为引物或被反转录成小片段的 DNA，干扰全长 cDNA 的合成，因而在反应中起副作用，应尽量去除它的活性。β 链具有以 RNA-DNA 杂交分子为底物的 5′→3′脱氧核酸外切酶活性。

反转录酶是一个具有多功能的酶，此外，尚有 DNA 依赖的 DNA 聚合酶活性、DNA 解旋酶活性。

二、反转录酶催化 DNA 合成的分子机制

反转录酶具有 5′→3′方向的聚合活性。该酶能利用具有 3′端 poly(A)的 mRNA 为模板，在引物 oligo$(dT)_n$(寡聚脱氧胸腺嘧啶核苷)的引导下，合成与 mRNA 互补的 DNA 链(即 cDNA，complementary DNA)，此 cDNA 称为第一链 cDNA。该反应的产物是一种 RNA-DNA 杂交分子，通过碱处理去处 mRNA 模板，再以第一链 cDNA 为模板，在特异性引物指引下，由大肠杆菌聚合酶 Ⅰ(或 *Taq* DNA 聚合酶)催化合成第二链 cDNA(图 4-2)。由于单链 cDNA 3′末端易于自身环化，形成发夹结构，因此需用 S1 核酸酶降解发夹结构。此种体外以 RNA 为模板合成 DNA 大量进行扩增的方法又叫做反转录 PCR(RT-PCR，reverse transcript polymerase chain reaction)。

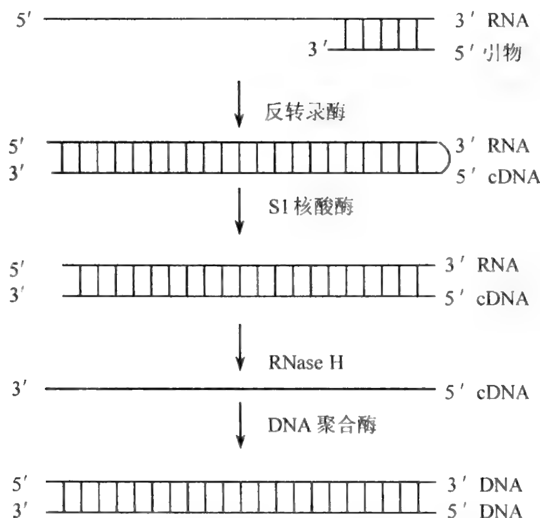


图 4-2　反转录酶催化 DNA 合成的分子机制

三、反转录酶在基因工程中的应用

真核生物基因组DNA十分庞大，且含有大量的重复序列。无论是采用电泳分离技术还是通过杂交的方法，均不能直接分离到目的基因片段。反转录酶的发现和应用，使通过mRNA得到DNA进行克隆成为现实。

1. 构建cDNA基因文库　在基因工程中，反转录酶的主要用途是转录mRNA为cDNA，制备基因片段，从而构建cDNA基因文库(详见第六章)。

2. 以ssDNA或RNA为模板制备核酸探针　以poly(A)mRNA为模板，在引物oligo $(dT)_n$ 的引导下，当存在四种dNTP时，反转录酶可催化cDNA的合成，若在dNTP中加入 α-^{32}P标记的dNTP(α-^{32}P-dNTP)时，则新合成的cDNA就成为放射性核素标记的DNA链，通过纯化，就可以获得单链^{32}P标记的cDNA探针。反转录酶的引物也可以是特异的寡核苷酸引物，还可以采用随机寡核苷酸作为引物(详见第十章)。

第四节　DNA聚合酶

大肠杆菌DNA聚合酶、大肠杆菌DNA聚合酶Ⅰ的Klenow大片段酶(Klenow酶)、T4 DNA聚合酶、T7 DNA聚合酶、反转录酶及 *Taq* DNA聚合酶等(表4-3)均是基因工程中常用的DNA聚合酶，其共同特点是，能够把脱氧核糖核苷酸连续地加到dsDNA分子引物链的3′-OH端，催化DNA分子聚合，而引物不从DNA模板上解离。

表4-3　DNA聚合酶特性

聚合酶名称	来　源	3′→5′外切活性	5′→3′外切活性	聚合反应速率	持续合成能力
E.coli DNA聚合酶	大肠杆菌	低	有	中	低
Klenow大片段酶	大肠杆菌	低	无	中	低
T4 DNA聚合酶	T4噬菌体	高	无	中	低
T7 DNA聚合酶	T7噬菌体	高	无	快	高
反转录酶	肿瘤病毒	无	无	低	中
Taq DNA聚合酶	栖热水生菌	无	有	快	高

一、大肠杆菌DNA聚合酶Ⅰ及应用

大肠杆菌DNA聚合酶(*E.coli* DNA polymerase)有三种不同的类型，即DNA聚合酶Ⅰ、DNA聚合酶Ⅱ、DNA聚合酶Ⅲ，分别简称为polⅠ、polⅡ和polⅢ。polⅠ和polⅡ的主要功能是参与DNA的修复过程，polⅢ同DNA的复制有关。在基因工程中，polⅠ与分子克隆的关系最为密切。

（一）DNA 聚合酶Ⅰ的特性

DNA 聚合酶Ⅰ是 1957 年由 Arthur Kornberg 首先发现的 DNA 聚合酶，又称 Kornberg 酶。此酶研究的最清楚而且代表了其他 DNA 聚合酶的基本特点。

1. DNA 聚合酶Ⅰ的生物学特性　由大肠杆菌 *pol*A 基因编码，是一种单链多肽蛋白质，相对分子质量为 109×10^3。此酶是一个多功能酶，具有 5′→3′的聚合酶活性、5′→3′核酸外切酶活性及 3′→5′核酸外切酶活性。酶的活性发挥需要 Zn^{2+}，在 37℃条件下，每个酶分子每分钟能催化 667 个核苷酸掺入正在延长的 DNA 链中。

2. DNA 聚合酶Ⅰ的聚合反应特点

（1）DNA 聚合酶Ⅰ必须在具有全部四种脱氧核苷三磷酸 dNTPs(dATP、dGTP、dCTP、dTTP)和 Mg^{2+} 存在时，才能催化合成 DNA 的互补链。

（2）催化 DNA 链的合成是按 5′→3′方向生长的，即在生长链的 3′-OH 末端掺入新的核苷酸，这就要求引物链带有 3′-OH。

（3）模板可以是 dsDNA，也可以是 ssDNA。

3. DNA 聚合酶Ⅰ的 5′→3′核酸外切酶活性——切除修复作用　5′→3′外切酶活性就是从 5′→3′方向水解 DNA 生长链前方的 DNA 链，即从游离的 5′-P 末端水解 DNA 分子，释放的产物主要是 5′-磷酸核苷，也可以有少量较大的核苷酸片段。这种酶活性只对 DNA 上配对部分（双链）的磷酸二酯键有切割作用，每次能切除 10 个核苷酸，在 DNA 损伤的修复中可能起着重要作用。

4. DNA 聚合酶Ⅰ的 3′→5′核酸外切酶活性——校对作用　从 DNA 链的 3′-OH 末端开始向 5′的方向水解 DNA 分子，主要功能是识别和切除不配对的 DNA 生长链末端的核苷酸，对维持 DNA 复制的真实性十分重要。当反应体系中没有反应底物 dNTP 时，由于没有聚合作用而出现暂时的游离现象，从而被 3′→5′外切酶活性所降解。若加入的核苷酸与模板不互补而游离时则被 3′→5′外切酶切除，以便重新在这个位置上聚合对应的核苷酸。只有在加入与模板互补的核苷酸时，才会使外切酶活性受到抑制，并继续进行 DNA 的合成。因此，3′→5′外切酶活性的主要功能是校对作用。

（二）DNA 聚合酶Ⅰ在基因工程中的应用

DNA polⅠ的 DNA 聚合酶活性和 5′→3′外切酶活性协同作用，可以使 DNA 链上的切口向前推进，即没有新的 DNA 合成，只有核苷酸的交换。这种反应叫缺口平移（nick translation）。当双链 DNA 上某个磷酸二酯键断裂产生切口时，DNA polⅠ能从切口开始合成新的 DNA 链，同时切除原来的旧链。这样，从切口开始合成了一条与被取代的旧链完全相同的新链。如果新掺入的脱氧核苷三磷酸为 α-^{32}P-dNTP，则重新合成的新链即为带有放射性核素标记的 DNA 分子，可以制备探针进行分子杂交实验（详见第十章）。

二、大肠杆菌 DNA 聚合酶Ⅰ的 Klenow 大片段酶

大肠杆菌 DNA 聚合酶Ⅰ全酶，经舒替兰酶（枯草杆菌蛋白酶）分解后，形成两个片段，

一个片段相对分子质量为 36×10^3，具有 5′→3′方向的核酸外切酶活性；另一个是相对分子质量为 76×10^3的大片段，具有 5′→3′聚合酶与 3′→5′方向的核酸外切酶活性。此大片段叫做大肠杆菌 DNA 聚合酶Ⅰ的 Klenow 片段（*E.coli* DNA polⅠ Klenow fragment），又称之为 Klenow 聚合酶或 Klenow 大片段酶。

在基因工程中，Klenow 大片段酶的主要用途有：

（1）修补经限制性内切酶消化 DNA 所形成的 3′隐蔽末端；

（2）标记 DNA 片段的末端；

（3）cDNA 克隆中第二链 cDNA 的合成；

（4）DNA 序列测定；

（5）随机引物末端标记法标记 DNA 的 3′末端，以制备基因探针（详见第十章）。

三、耐热 DNA 多聚酶

耐热 DNA 多聚酶又叫 *Taq* DNA 聚合酶（*Taq* DNA polymerase），最初由 H A Erlich 于 1969 年从栖生水热菌（*Thermus aquaticus*）YT-1 菌株中分离纯化出。此菌生活在 70～75℃的环境中，在嗜热脂肪芽孢杆菌及其他嗜热菌中也得到类似的耐热 DNA 聚合酶。

Taq DNA 聚合酶基因全长 2496 个碱基，编码 832 个氨基酸，相对分子质量为 94×10^3。该酶最为突出的特点是具有良好的热稳定性。95℃保温 2 小时，酶的活性仍保持良好。该酶在 70～75℃时具有最高的生物学活性，能以从高温变性的靶 DNA 分子中分离出的 ssDNA为模板，从分别结合在扩增区段两端的引物为起点，按 5′→3′方向合成互补 DNA 的新生链。75～80℃条件下，每一个酶分子每秒可延伸约 150 个核苷酸，温度降低时，*Taq* DNA 聚合酶催化新生链延伸的速度明显下降。该酶是 PCR 反应中重要的酶，PCR 过程中，解链温度多采用 94℃、30 秒，循环约 30～35 次，*Taq* 酶完全能保证 PCR 反应的正常进行而自身不会因高温而失活（详见第九章）。*Taq* DNA 聚合酶的发现，使 PCR 技术得到迅速发展和广泛应用。

Taq DNA 聚合酶除具有 5′→3′聚合酶活性外，尚具有 5′→3′外切酶活性，但缺乏3′→5′外切酶活性。因而，在 PCR 反应中如果发生单核苷酸的错配，不能够及时校正，一般情况下出现碱基错配的几率为 2.1×10^{-4}。*Taq* DNA 聚合酶活性发挥对 Mg^{2+} 浓度非常敏感。Mg^{2+} 浓度高时，酶活性易受到抑制，一般在不同的反应体系中应适当进行调整。

四、反 转 录 酶

为基因工程中最重要的核酸酶之一（详见本章第三节）。

五、T4 DNA 聚合酶及 T7 DNA 聚合酶

T4 DNA 聚合酶（T4 DNA pol）是从 T4 噬菌体感染的大肠杆菌培养物中纯化出的聚合酶，由噬菌体基因 *43* 编码，为单链多肽酶，具有 5′→3′聚合酶与 3′→5′方向的核酸外切酶活

性。在催化DNA合成时，可以单链DNA作模板，也可以有缺口的双链DNA作模板，但后者需要有基因32蛋白的辅助。利用单链DNA为模板进行扩增，可同时利用它作为引物，即此单链DNA的3′端能环绕其本身的某一顺序形成氢键配对，3′端的未杂交部分即被T4 DNA聚合酶的3′→5′外切酶活性切去，然后在其作用下从3′-OH端开始聚合，合成该模板DNA的互补链，再以互补链为模板合成原来的单链DNA。此酶的功能是可用以催化标记DNA平末端或隐蔽的3′末端。

T7 DNA聚合酶是从T7噬菌体感染的大肠杆菌宿主细胞中纯化出的一种核酸酶。具有DNA聚合酶活性及3′→5′方向的核酸外切酶活性。

六、真核生物的DNA聚合酶

真核生物中也具有DNA聚合酶，但这些聚合酶多数都没有3′→5′或5′→3′外切酶活性。其聚合反应机制与原核生物的聚合一样。主要有：①DNA聚合酶α，占酶总量的80%～90%，相对分子质量为300×10³，含有4个或5个亚基，主要负责染色体DNA的复制。②DNA聚合酶β，相对分子质量为45×10³，仅含有一条链，主要作用是修复核内DNA。③DNA聚合酶γ，相对分子质量为140×10³，存在于线粒体内，负责催化线粒体DNA的复制。最近从兔的骨髓细胞质中分离到一种新的DNA聚合酶，这种酶与上述真核生物的DNA聚合酶不同，而与原核生物的聚合酶类似，具有3′→5′核酸外切酶活性，称为DNA聚合酶δ。另外，从酵母、原虫等低等真核生物中分离到一些DNA聚合酶。一般来说，这些低等生物都没有相对分子质量低的DNA聚合酶β，而且与原核生物的DNA聚合酶相似，有3′→5′外切酶活性。

第五节　碱性磷酸酶

有两种不同来源的碱性磷酸酶，一种是从大肠杆菌中分离纯化出的，叫做细菌碱性磷酸酶(bacterial alkaline phosphatase，BAP)；另一种是从小牛肠中分离纯化出的，叫做小牛肠碱性磷酸酶(calf intestinal alkaline phosphatase，CAP)。其共同特性是，特异性切去DNA或RNA分子的5′-P，从而使DNA或RNA分子的5′-P末端转换为5′-OH末端，即所谓的核酸分子脱磷酸作用。碱性磷酸酶的底物可以是ssDNA或dsDNA或RNA，也可以是核糖或脱氧核糖核苷二磷酸。



$$5'\text{P-DNA} \xrightarrow{\text{BAP}} 5'\text{OH-DNA} + \text{Pi}$$

$$5'\text{P-RNA} \xrightarrow{\text{BAP}} 5'\text{OH-RNA} + \text{Pi}$$

CAP和BAP在理化特性上稍有不同，CAP的活性比BAP高10～20倍，且CAP对热较不稳定，在SDS中加热到68℃即可完全失活，而BAP具有热抗性，终止它的作用较为困难，需要酚/氯仿反复抽提多次，因而BAP不如CAP应用广泛。

碱性磷酸酶在基因工程中的应用主要有：

1. 制备5′末端标记的核酸探针　碱性磷酸酶脱磷酸作用的产物是具有5′-OH末端的

核酸分子，有底物[γ-^{32}P]ATP 存在的条件下，在 T4 多核苷酸激酶催化下，将 ATP 分子上的 γ-^{32}P 基团转移到 DNA 或 RNA 分子的 5′-OH 上，就可以获得 5′末端标记的核酸探针（详见第十章）。

2. 防止 DNA 重组中载体的自身环化　在体外 DNA 的重组中，为防止线性化的载体分子自我环化或自身连接，在连接反应之前，用碱性磷酸酶处理载体分子，除去 5′-P 基团，加入目的 DNA 进行重组，再对此反应混合物进行热处理，则没有与目的 DNA 重组的载体分子解链，恢复为线性构型，这样就不能够形成转化子克隆，而只有载体分子与目的 DNA 形成的重组体才可以在受体细胞内形成阳性克隆，从而很容易筛选出来。

第六节　单链核酸内切酶——S1 核酸酶

S1 核酸酶是一种从稻谷曲霉（*Aspergillus oryzae*）中分离纯化出的、高度单链特异的核酸内切酶，其主要的功能是降解 ssDNA 或 RNA，形成具有 5′-P 末端的单核苷酸或寡核苷酸片段。

S1 核酸酶活性的发挥需要 Zn^{2+} 作辅助因子，最适 pH 为 4.0～4.5，当 pH 上升到 4.9 时，降解速率则会下降 50%。对 NaCl 的浓度要求的范围可以变动较大，为 10～300 mmol/L，但最佳浓度为 100 mmol/L。EDTA、枸橼酸、磷酸盐缓冲液、SDS 等，都可以抑制 S1 核酸酶的活性。

S1 核酸酶在基因工程中的应用主要包括：

1. RNA 分子定位　给 RNA 分子定位是 S1 核酸酶在分子生物学研究中的一个最主要的功能。一个 RNA 分子是由其 DNA 模板的一定的核苷酸序列编码的，若这个 RNA 分子同包括该模板的 DNA 分子（碱基组成多于该模板）杂交，则形成具有单核苷酸链的 RNA-DNA 杂种双链分子，利用 S1 核酸酶去除单链尾巴，回收双链分子，碱变性后进行琼脂糖凝胶电泳或聚丙烯酰胺凝胶电泳，就可以测出 RNA-DNA 杂种分子中的 DNA 片段的大小，从而对 RNA 分子进行定位。

2. 测定真核基因中内含子的位置　由两外显子拼接的 mRNA 与模板 DNA（由两个外显子和外显子之间的内含子组成）杂交，形成三种 RNA-DNA 杂种分子：一种是组成的 RNA-DNA 杂种分子在两外显子之间形成一个单链 DNA 环，利用 S1 核酸酶降解掉单链 DNA 环和两端的单链 DNA 尾巴，可获得两外显子全长的 DNA 分子；另两种是各由一个外显子 DNA 与 RNA 组成的 RNA-DNA 杂种分子，S1 核酸酶降解掉多余的内含子和两端的单链 DNA 尾巴，可获得与每个外显子相对应的 DNA 分子。通过琼脂糖凝胶电泳，可得到各个 DNA 的相应谱带，从而确定内含子的位置。

此外，S1 核酸酶可以消化限制性核酸酶产生的黏末端，形成平齐末端，降解 cDNA 合成过程中形成的发卡环结构，在质粒组建和基因重组中被广泛应用。

第七节　末端脱氧核苷酸转移酶

末端脱氧核苷酸转移酶（terminal deoxynucleotidyl transferase），简称末端转移酶（termi-

nal transferase)，是从小牛胸腺中纯化出的一种相对分子质量较小(34×10^3)的碱性蛋白酶。

一、末端转移酶的生物学功能

催化5′脱氧核苷三磷酸转移到另一个DNA分子的3′-OH末端，即进行5′→3′方向聚合作用。与DNA聚合酶不同，末端转移酶不需要模板的存在，就可以催化DNA分子聚合。末端转移酶的模板是具有3′-OH末端的ssDNA或具有延伸3′-OH末端的dsDNA。

二、末端转移酶在基因工程中的应用

1. 给DNA片段3′-OH末端加尾　末端转移酶是一种非特异性酶，当反应混合物中只有一种dNTP时，形成只有一种核苷酸组成的3′尾巴，称之为同聚物尾巴(homopolymeric tail)。在基因工程中，利用末端转移酶分别给载体分子和外源DNA片段加上互补的同聚物尾巴，造成人工黏末端，有利于DNA重组(详见第六章)。

2. 催化DNA片段的3′-OH末端标记，用于DNA序列分析　在DNA序列分析中，末端转移酶催化[α-^{32}P]-3′脱氧核苷酸加入到DNA分子的3′末端，使该DNA不再具有游离的3′-OH末端，可终止单核苷酸的掺入作用，通过凝胶电泳后的放射自显影，就可以读出DNA的序列。

3. 制备3′末端标记的DNA探针　末端转移酶可催化标记的dNTP掺入到DNA片段的3′-OH末端，制备3′末端标记的探针(详见第十章)。

表4-4总结了几种重要工具酶的酶学性质及其应用，供参考。

表4-4　几种重要工具酶的酶学性质及其应用

酶	来　源	活　性	底　物	辅助因子	特　点	基因工程中的主要用途
限制酶(Ⅱ型)	微生物	内切	dsDNA	Mg^{2+}	特异性识别与切割，产生平齐或黏末端	①基因重组 ②组建物理图谱 ③制备探针 ④DNA测序
连接酶	*E.coli* T4噬菌体	连接两DNA的3′-OH与5′-P 5′→3′聚合	dsDNA ssDNA	Mg^{2+} ATP	对黏末端的连接效率高于平末端	①DNA重组 ②构建新质粒
DNA pol Ⅰ	*E.coli*	5′→3′外切 3′→5′外切	dsDNA dsDNA	Mg^{2+}	对dsDNA的外切被5′→3′聚合活性抑制	①制备探针 ②DNA测序
反转录酶	RNA肿瘤病毒	RNA→DNA合成 DNA→RNA合成 RNaseH 解旋	ssRNA ssDNA RNA-DNA cccDNA	Mg^{2+}	RNA指导 DNA指导 RNA-DNA杂交链 解超螺旋	①制备cDNA，构建基因文库 ②制备探针

续表

酶	来源	活性	底物	辅助因子	特点	基因工程中的主要用途
S1核酸酶	稻米曲霉	降解单链核酸	ssRNA ssDNA	Zn^{2+}		①RNA定位 ②测定间隔子 ③构建新质粒
末端转移酶	小牛胸腺	在3′末端掺入核苷酸	ssDNA	Mg^{2+}	Co^{2+}存在时，可用dsDNA为模板	①3′同聚尾 ②人工黏末端 ③构建新质粒
*Bal*31	乳白短杆菌	外切	dsDNA	Ca^{2+}	5′、3′两端同时等速外切	测定DNA中限制酶识别位点
BAP CAP	细菌 小牛肠	切去5′磷酸	ds/ssDNA	Mg^{2+}		①防止载体或片段自身环化 ②制备探针

（刘素侠）

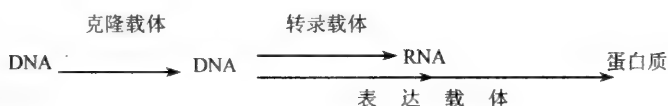
第五章　基因工程的克隆载体

Chapter 5. Genetic Cloning Vector

第一节　概　　论

基因工程的目的和基本手段，是选用合适的载体把供体 DNA(外源基因)运载到相应的受体细胞内，从而复制、扩增出大量的目的 DNA 分子或者转录、表达为相应产物，所有这些都必须选择合适的载体。

基因工程中所使用的载体可分为：克隆载体(cloning vector)、转录载体(transcription vector)和表达载体(expression vector)。克隆载体是指可以携带外源基因或 DNA 片段进入受体细胞并使外源 DNA 大量扩增的复制单元，适用于复制扩增外源 DNA。转录载体是以 DNA 为模板、用 RNA 聚合酶Ⅱ合成 RNA 的过程，多采用体外转录体系。表达载体是用于表达原核或真核细胞中某个特定基因所编码的蛋白质产物的载体，其克隆位点的上游含有强启动子、下游含有转录终止序列，适用于表达某特定蛋白质。克隆、表达的基本流程是：



原核与真核细胞的基因结构及其表达、调控有不同的特点。原核细胞基因组的主要特点是：①染色质为环状双股 DNA 分子，外无核膜与胞质隔开，使转录与翻译偶联、连续进行；②具有操纵子结构，多个基因共受同一调节基因调控，这能使插入到该多顺反子区域的外源 DNA 同样被复制；③结构基因多为单拷贝(即基因不重复)；④特定区域分布着特异性的 DNA 顺序(元件)：如调节基因、控制单元和结构基因，转录的起动区、终止区等。这些特异的 DNA 元件能精确、高效地调控着 DNA 的复制、RNA 的转录(图 5-1)。原核细胞基因组的这些特点足以把一个外源 DNA 分子连接、插入到原核细胞 DNA 复制体系的特定区段而被扩增；同样，亦能把 DNA 复制体系中的某结构元件进行改造，引入到其他载体系统中，构建成新的载体。

基因工程中的克隆载体应具有如下特性(图 5-2)：①具有复制子，能在宿主细胞内自主地复制，并携带重组 DNA 分子一同扩增；②有单一限制性内切酶的酶切位点或多克隆位点(multiple cloning site，MCS)，MCS 是人工合成的一段含有多个不同的单一限制性内切酶切点的 DNA 序列，它有利于外源基因插入该区域；③有选择性遗传标记，如抗药基因、酶基因、营养缺陷体及噬菌斑形成能力等，便于筛选出阳性重组体；④拷贝数高(10～200 个/每个细胞)，易于分离；⑤生物安全性好。

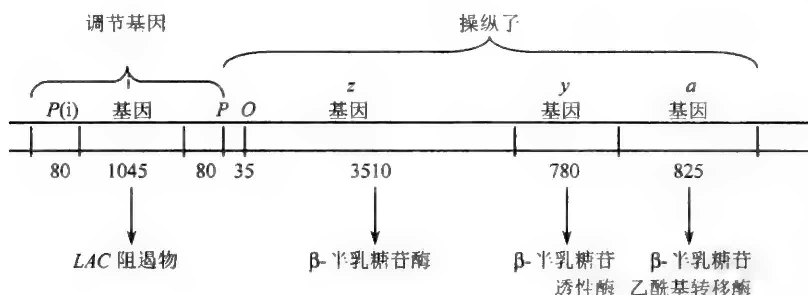


图 5-1 原核细胞基因组的结构特征——*Lac* 操纵子结构模型

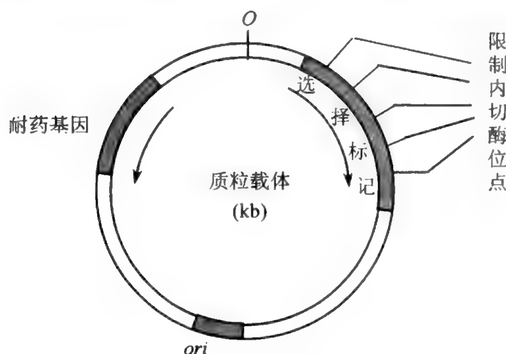


图 5-2 克隆载体模式图

目前所使用的载体，按其特性可分为六类：①质粒（plasmid）；②λ噬菌体(bacteriophageλ)；③黏性质粒(cosmid)；④M13 噬菌体(phage M13)；⑤酵母(yeast)载体；⑥真核细胞病毒载体(eukaryotic virus vector)。前四类克隆载体用于原核细胞中；酵母和病毒载体则属于真核细胞克隆或表达载体。

第二节 细菌质粒的生物学特性

一、质粒的定义

质粒是存在于细菌等微生物细胞染色质以外的共价闭环的双股 DNA 分子，具有独立自主复制和调控的能力，可赋予宿主细胞一定的生物学性状。

二、质粒的生物学特性

1. 大小 根据质粒的大小，分为 2 型：①小型质粒：约 1.5～15kb 长，易于转化，因不带可传递基因(*tra*)，故不能接合传递；②大型质粒：约 60～120kb，带有 *tra* 基因，能接合传递。因太大，不易转化。

2. 复制类型 根据质粒的复制型式，分 2 类：①严紧型质粒(strengent plasmid)：大型质

粒多属之,它拷贝数少(1～2个/每个细胞),具自身传递能力。其DNA复制依赖宿主细胞的DNA聚合酶Ⅲ及蛋白质合成体系(即酶系统两者相同),与宿主细胞染色体DNA的复制相偶联且呈双向复制,故复制受宿主细胞的严格控制;②松弛型质粒(relaxed plasmid):多半为小型质粒,拷贝数多(10～200个/每个细胞)。其DNA复制是在宿主细胞松弛控制下进行的,只需DNA聚合酶Ⅰ参与,与染色体复制不同步,适用于基因工程中做质粒载体。

用氯霉素可扩增松弛型质粒DNA,因氯霉素可与细菌的50S核糖体亚基结合,抑制肽基转移酶活性,从而抑制DNA聚合酶、蛋白质的合成,使细菌DNA和严紧型质粒DNA的复制停止。松弛型质粒的DNA复制可依然进行,拷贝数大量扩增(1000～3000个),因此用氯霉素可大量扩增质粒DNA。

3. 拷贝数　拷贝数是指在宿主细胞分裂前质粒DNA自我复制的份数。小型质粒拷贝数多,易改造为克隆载体;大型质粒拷贝数少,不易改造。

4. DNA构型　电镜下观察DNA有3种构型:①超螺旋型(Ⅰ型即CCC型,covalently closed circle DNA或SC型,supercoiled circle):因分子结构为共价闭环的双股DNA分子,呈超螺旋结构。②开环型(Ⅱ型即OC型,open circle):双股DNA有一股断开,形成“缺口”。③线型(Ⅲ型即LC型,linear circle):DNA双股在某处完全断开。

5. 电泳特点　上述3种构型的DNA分子,在琼脂糖凝胶中有不同的泳动速度,其泳动速度主要取决于凝胶的浓度,也受电流、缓冲液、离子强度和溴化乙锭(EB)的浓度影响。电泳迁移率的大小一般是:CCC型>LC型>OC型,借此可分析DNA分子的构型,同在LC型条件下可测量DNA分子的大小。

6. 不相容性　利用同一复制系统的不同质粒不能稳定地共存同一宿主细胞,称为不相容性(incompatibility),如pBR322(图5-3,图5-4)和pUC系列,均带pMB1复制子而彼此不相容,但可与带pSC101质粒的细胞相容。带有不可互换成分的复制子的质粒可共存同一细胞,称之为相容性(compatibility),基因操作应考虑到复制子等因素是否具相容性。

三、质粒的分类

1. 按复制型式　分为严紧型和松弛型复制。

2. 按基因转移性　分两类:①传递性质粒:常为大型、严紧型质粒;②非传递性质粒:多为小型、松弛型质粒。

3. 按遗传性状、产物分类　①抗生素抗性:如氨苄西林、氯霉素抗性。②限制酶-修饰酶(R-M)系统:如hsdR$^-$菌株可使DNA甲基化,但不限制外源DNA,即(r_k^- m_k^+);hsd$^-$菌株为限制和甲基化双缺陷型(r_k^- m_k^-)。

第三节　质粒载体

一、概　　述

天然的细菌质粒并不能完全满足基因工程所要求的条件,必须经过改造才能成为理想、

实用的载体。改造的方法是基因突变、修饰、重排、重组等。

改造载体的目的是调整载体的结构、提高转化率。载体改造的常用手段包括:①减小相对分子质量。一般载体越小,转化率愈高。②在载体上引入"多克隆位点"(MCS,multiple cloning site),便于外源 DNA 片段插入该区。③在质粒中引入特定用途的辅助序列,使载体功能更加完善和专业化。例如:引入 *lacZ'* 基因用于蓝白斑筛选;加入病毒启动子和其他元件在真核细胞表达;加入真核和原核表达元件,可做成穿梭质粒载体或者真核细胞的克隆、表达载体。

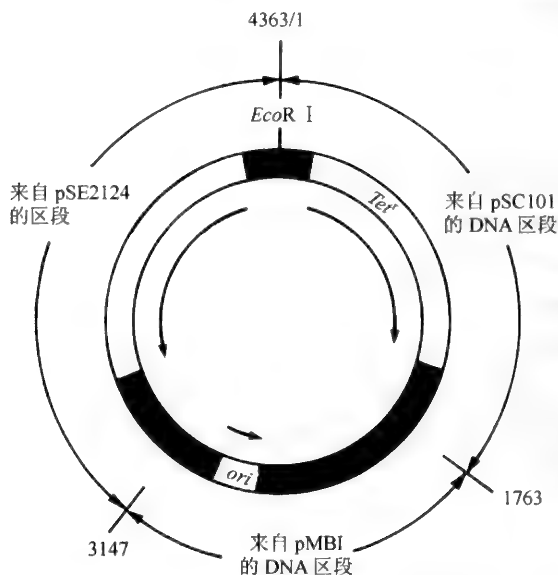


图 5-3　pBR322 质粒载体的结构来源图

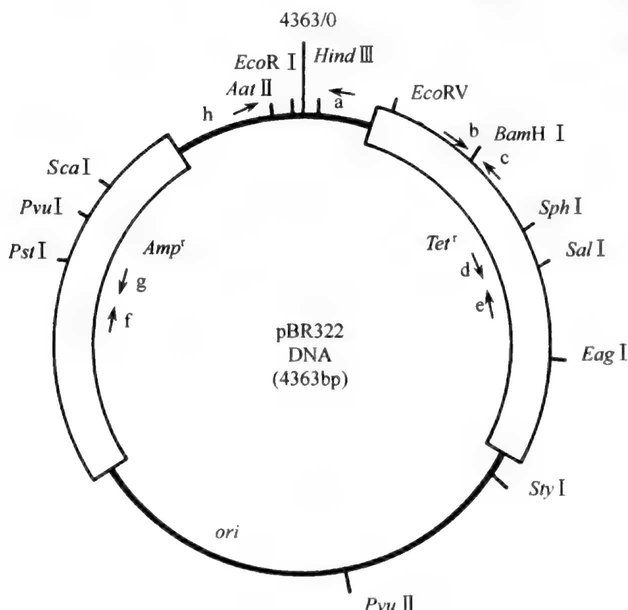


图 5-4　pBR322 物理图谱

二、常用的质粒克隆载体

1. pBR322 质粒

(1) pBR322 质粒的构建：pBR322 来源于三个天然质粒 pSC101、ColE1(带有 pMB1 复制子)和 pSF2124(表 5-1)，经过融合→转位→重排→缺失，最终组建成质粒 pBR322(图 5-3、4)。

表 5-1 三个天然质粒的基本特性

质 粒	复制型	相对分子质量/($\times 10^6$)	单一酶切点	选择标记
pSC101	严紧型	5.8	*Eco*RI	Tc^r
ColE1	松弛型	4.2	*Eco*RI	Col^{imm}
pSF2124	松弛型	7.4	*Eco*RI、*Bam*HI	Ap^r

(2) pBR322 的物理图谱与识读(Bolivar et al. 1977)：DNA 物理图谱(physical map)，是指在 DNA 分子上标明限制性内切酶存在的位点、数目、排列顺序及特征性功能标记的图形，亦称限制图谱(restriction map)。

物理图谱的表现形式分为三类：①环形；②线形(含 DNA 一级序列)；③复合型(环形与线形相结合)。

pBR322 物理图谱见图 5-4，识读该图可获得如下信息：① DNA 大小(4363bp，2.6×10^6)；②零点(环形图谱的起止点)；③*ori*(复制起始点，箭头示复制方向)；④选择标记(两个抗性标记：Ap^r、Tet^r；箭头表示转录方向)；⑤重要的单一酶切点(分别密集分布在 Ap^r 区、Tet^r 区，提供了插入灭活和选择标记)；⑥内切酶旁的数码，表示该酶的切口位置。

(3) pBR322 的应用：①常规的原核细胞克隆载体(用于扩增、保存 DNA，片段的长度多小于 2kb)，应用最广。②改造成衍化质粒，如 pBR325、pBR327、plink322 等。③利用 pBR322 构建原核表达载体，包括非融合性表达载体 pKK223-3、分泌型表达载体 pINⅢ系统、及融合表达载体 pGEX 系统。

(4) pBR322 的宿主细胞：包括 HB101、JM107、JM109、LE392 等。这些宿主细胞的共同特征主要是：其 F′因子上的 $lacI^q$、$lac\Delta M15$ 基因经过修饰，致 α-肽表达缺陷。

(5) 基因克隆选择受体细胞的原则：①与所选择载体相匹配。②载体-受体细胞之间是否存在 lacZ 表达-缺陷互补、药物抗性标志，便于筛选。③宿主细胞相容性好。

2. pUC 载体

(1) pUC 载体系列：包括 pUC8/9、pUC12/13 和 pUC18/19，成对排列，为 Mssin 等人 1983、1985 年先后构建。

(2) pUC18/19 的特点(图 5-5A 限制图谱)：pUC18/19 是由 pBR322 和 M13 噬菌体构建的双链 DNA 质粒载体，长 2.69kb。含有 pBR322 的复制起始点、大肠杆菌 *LacZ′* 基因、Ap^r 抗性基因；含有 M13 的 MCS(置于 *LacZ′*中)，便于插入目的 DNA。pUC 质粒系列是应用较普遍的质粒载体。

pUC 质粒常是成对的，成对载体的其他特性完全相同，只是 MCS 的排列方向互为相反(图 5-5B)。

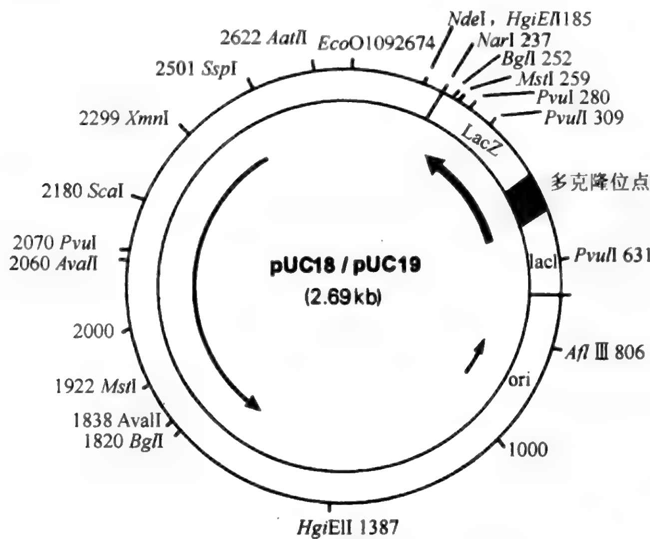


图 5-5A pUC18/19 克隆载体限制图谱

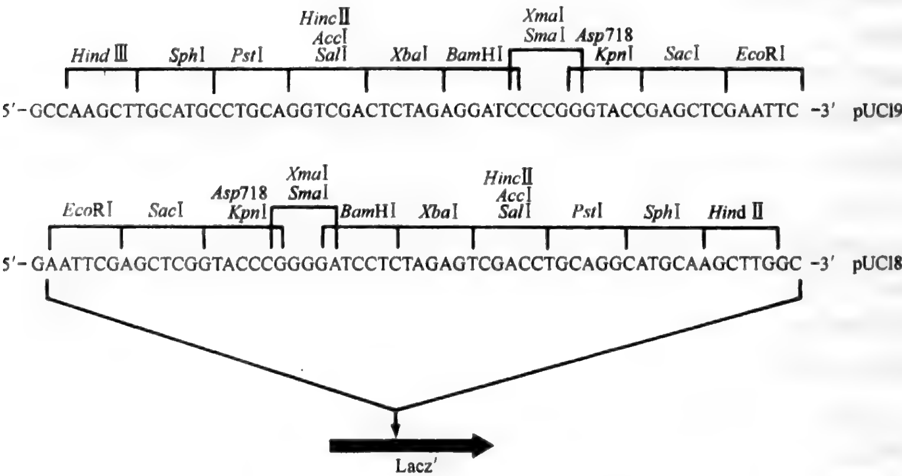


图 5-5B pUC18 和 pUC19 载体中的多克隆位点示意图

pUC 质粒呈多拷贝(500～700 个/每个细胞)对数生长,可用氯霉素扩增。

(3) LacZ′基因与蓝白斑筛选:*lacZ′*基因是大肠杆菌 *lacZ* 操纵子的 DNA 片段,它编码 β-半乳糖苷酶氨基端的 146 个氨基酸形成 α-肽,该 α-肽与宿主细胞中 F′因子上的 *lacZ′* $\Delta M15$ 基因(α-肽缺陷型)的产物互补,产生完整、有活性的 β-半乳糖苷酶。这种酶分解生色底物(X-gal,5-溴-4-氯-3-吲哚-β-D-半乳糖苷),形成蓝色菌落;当外源基因插入 MCS 后,*lacZ′*α-肽基因的读码框架被破坏,不能合成完整的 β-半乳糖苷酶分解底物 X-gal,菌落呈白色。用这种方法可筛选阳性重组子,并称为"蓝白斑"筛选。"蓝白斑"筛选法应用十分广泛。

(4) 受体细胞:包括 JM101、JM105、JM107 和 JM109 菌株等。这些菌株都是带琥珀突变载体的允许细胞,虽经甲基化修饰但并不限制转染的 DNA,可进行 β-半乳糖苷酶的 α-肽

互补。

(5) pUC 质粒载体的优点:①相对分子质量小(约 2.7kb)、拷贝数高。②具有多克隆位点,方便于克隆、转移、测序等。③可进行蓝白斑筛选,用组织化学法检测重组体。

3. TA 克隆载体(引自 invitrogen Clark,J M 1988;Mead,D et al.1991)

(1) 原理: TA 克隆技术是为了克隆 PCR 产物而设计的,它利用 *Taq* 聚合酶(和某些聚合酶的混合物)具有末端转移酶活性的特点,能把 PCR 产物的 3′末端加上一个非模板依赖的 A,所用 T 载体是带有 3′T 突出端的载体,在连接酶作用下可以高效(4 小时内)、一步到位地把 PCR 产物直接插入到该 T 载体中。

目前 TA 克隆载体已系列化、专业化、商品化,试剂品质优良、应用效果也好。TA 载体的突出特点是获取、连接外源 DNA 不受酶切点的限制,能直接克隆 PCR 产物,TA 克隆技术大大提高了 DNA 克隆的效率。近几年国内已广泛应用。

(2) 载体 pCR2.1 的主要特点(图 5-6):①3′-T 突出端能直接与 *Taq* 聚合酶扩增的 PCR 产物相连接;②体外 RNA 转录时选择 T7 启动子;③通用的 polylinkerT 接头两侧有 EcoRI 酶切位点,使容易切取插入子;④有 Kan^r 和 Amp^r 抗性选择标志;⑤可用蓝/白菌落

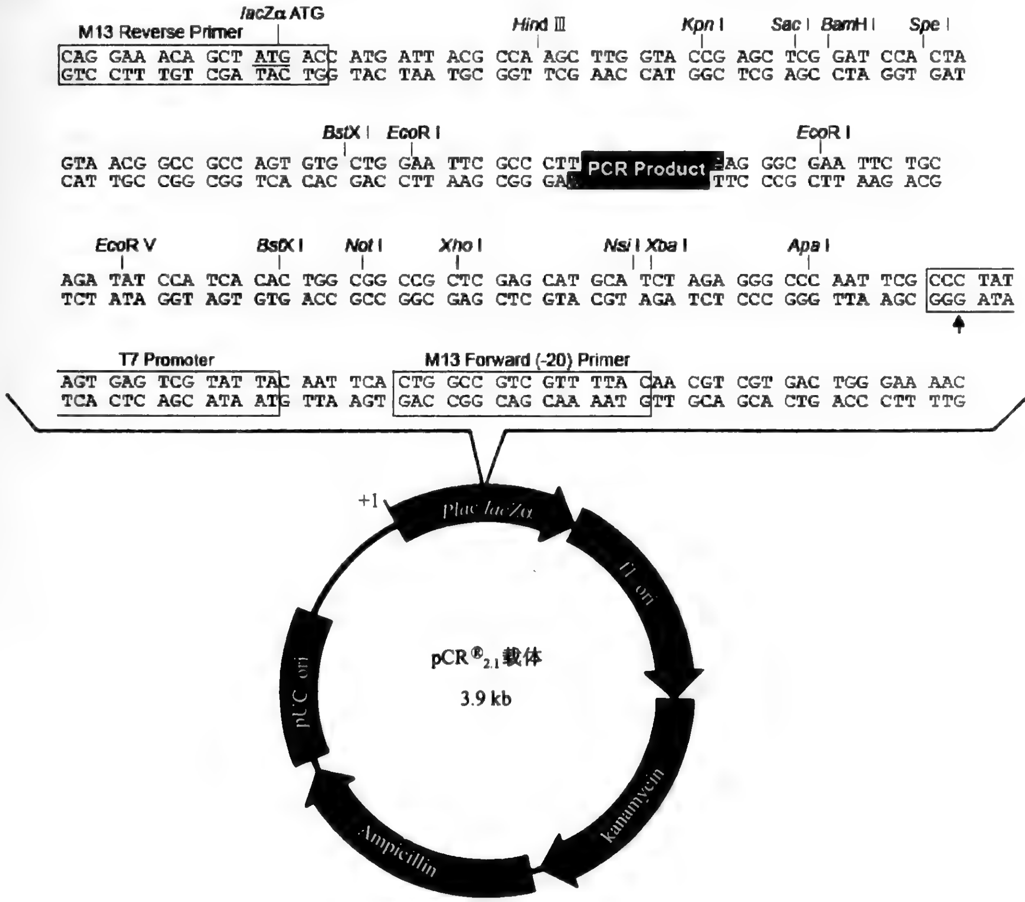


图 5-6 pCR2.1 克隆载体

筛选；⑥M13 提供了测序的正向和反向引物结合位点。

(3) TA 克隆方法的优点：①合成引物不必含有限制内切酶位点；②不需用 PCR 修饰；③减少 PCR 产物的纯化过程；④不需要连接的接头。

(4) 受体细胞的选择：为确保结果可靠应选择：*InvαF E. coli*（表型为：F′endA1 recA1 hsdR17(r_K, m_K^+) supE44 thi-1 gyrA96 relA1Φ80 lacZΔM15(lacZYA-argF)U169)。

该受体菌株是一个高拷贝数的质粒，其表型提供了以下特征：①用蓝白菌落可筛选出携带 β-半乳糖苷酶（lacZΔM15）的 α-肽的重组克隆；②能减少转化质粒的同源重组（recA）；③可提高 DNA 复制水平；④研究来自载体中携有 f1 复制起始点的单股 DNA。

(5) 应用：TA 载体为一克隆载体。主要应用于克隆 DNA，也可用于转录。

4. 其他质粒载体　从 pUC 系列派生了许多能在体外转录和克隆基因的载体，它们的共同特点是都带有噬菌体（T3、T7 或 SP6）编码的依赖 DNA 的 RNA 聚合酶转录单位的启动子，使载体能体外转录插入的外源 DNA，但是它们又各有特点。例如：pGEM3/4：它含有 *SP6*、*T7* 启动子和 *MCS*，便于插入、克隆外源 DNA。

pGEM-3Z/4Z：它比 pGEM-3/4 长出一段 *LacZ′* 序列，用蓝白斑筛选重组体（图 5-7）。

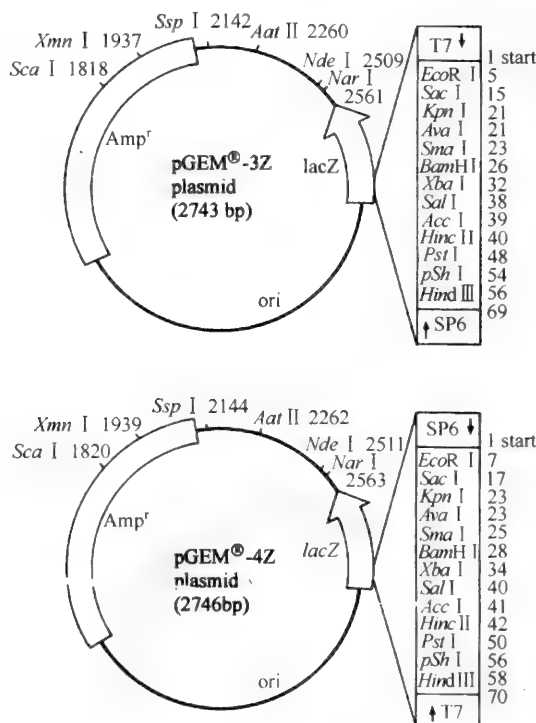


图 5-7　pGEM-3Z/4Z 的物理图谱图

pGEM-3Zf(+/-)：这是一对带有 f1 噬菌体 DNA 复制起点的载体，在该起点是以互为相反的方向在体内产生单链 DNA 或者在体外产生 RNA，这些产物能与插入多克隆位点的外源 DNA 双链中任意一条链互补（图 5-8）。pGEM-5Zf(+/-)和 pGEM-7Zf(+/-)与 pGEM-3Zf(+/-)的不同，是多克隆位点内限制性内切酶的数量和组成有所不同。

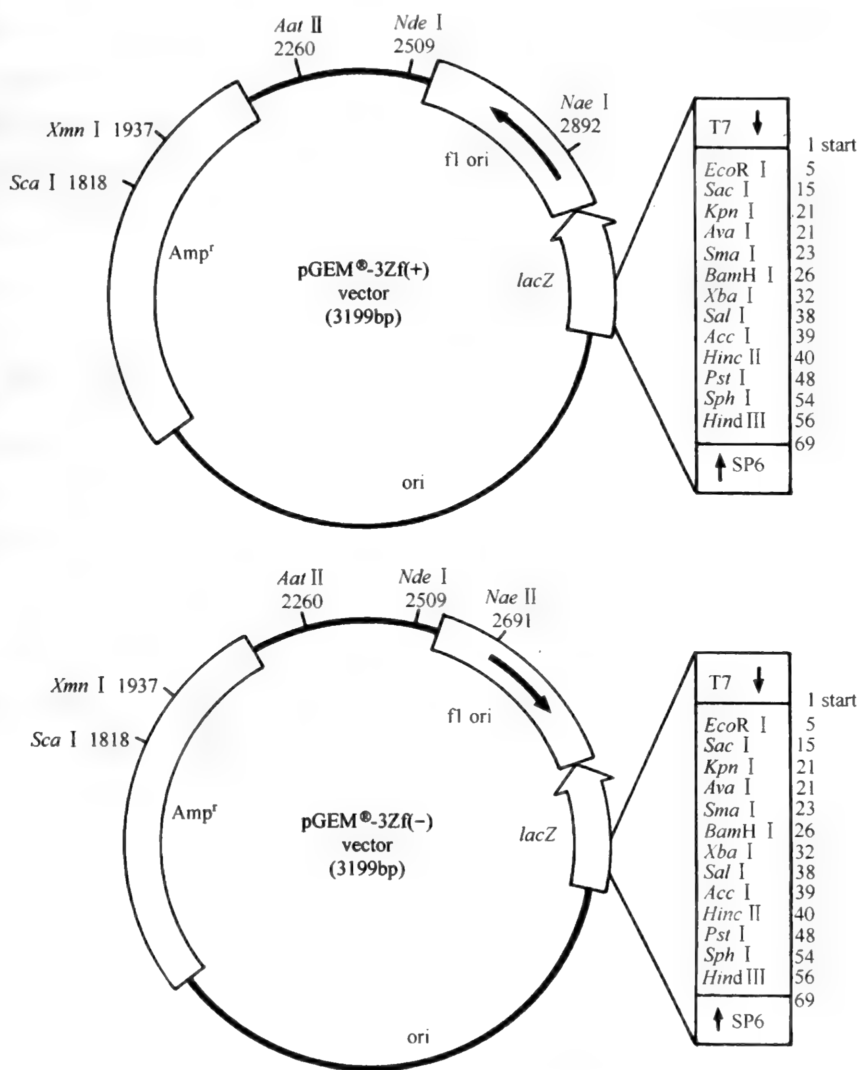


图 5-8　pGEM-3Zf(+/-)的物理图谱

第四节　λ噬菌体载体

一、λ噬菌体

(一) λ噬菌体的基本特性

λ噬菌体是大肠杆菌的温和型 DNA 噬菌体，属长尾噬菌体科。能自主复制，靠长尾末端吸附宿主细胞壁而感染。进入细胞内生长有两种形式：溶菌型(lytical pathway)和溶原型

(lysogenic pathway 或整合型)。溶菌生长时形成清亮的噬菌斑，易被识别，溶原生长时噬斑浑浊。λ 噬菌体基因组 DNA 近 50kb，中段有约 20kb 的区段对噬菌体的溶菌生长不是必需的，因此能被外源 DNA 取代，这种重组的 DNA 可被包装成具有感染性的噬菌体。在该 DNA 分子的两端各有 12 个碱基互补的 5′端单链，形成黏性末端位点(cos，cohensive-end site)，*cos* 位点的黏合、环化与 DNA 复制、转录以及感染有关。在感染早期环形 λDNA 按 θ 形式双向复制，感染晚期则以滚环方式复制出线性排列的多基因组连体的 DNA 分子，经切割、*cos* 环化、包装成熟。

(二) λ 噬菌体基因组和转录

1. λ 噬菌体基因组　λ 噬菌体基因组全长 48 502bp，为两端不闭合的线性双螺旋 DNA，由左、右两臂组成，编码 66 个基因，其中 46 个读框已经确定(图 5-9)。

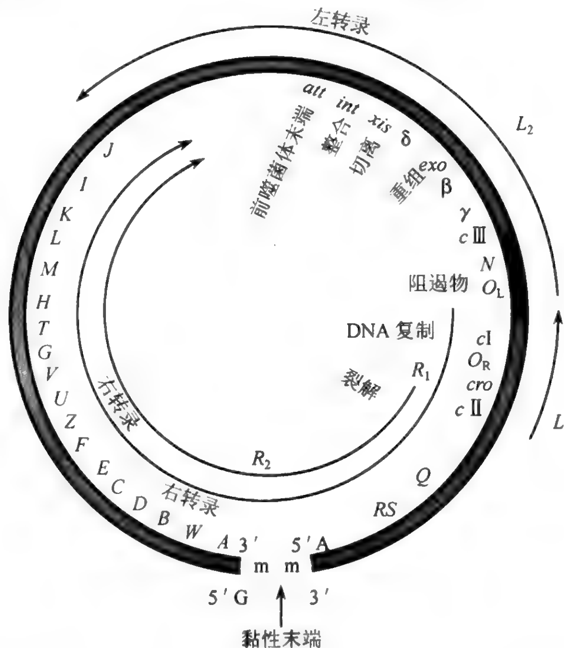


图 5-9　λ 噬菌体的遗传和分子图谱

互补 DNA 末段用 3′和 5′表示，基因用字母表示。*att*：噬菌体附着到宿主的位点；DNA 链向左转录(逆时针方向)，其左黏性末端(m)为 5′G，两个操纵子分别用 L_1 和 L_2 表示；DNA 链向右转录(顺时针方向)，其右黏性末端(m)为 5′A；两个操纵子用 R_1 和 R_2 表示。特别重要的基因：*c* Ⅰ：阻遏蛋白基因，O_R：R_1 转录的操纵基因，O_L：L_2 转录的操纵基因，*cro*：第二个阻遏蛋白基因，它阻遏 L_1、L_2 和 R_2 的转录；免疫区：*c* Ⅰ、*cro*、O_R 和 O_L 位于此；*N*：阻遏 cro 终止的调节基因；*J*～*U* 编码尾蛋白的基因；*Z*～*A* 编码头蛋白基因

λ 噬菌体的基因可分为控制复制的和转录调控的两类基因：

(1) 控制复制的基因及其功能简介

1) 头部基因：由 *A*，*W*，*B*，*C*，*D*，*E*，*F* 等 7 个基因组成，负责编码 λ 噬菌体头部蛋白质和具包装功能的蛋白质。

2) 尾部基因：由 Z,U,V,G,T,H,M,L,K,I,J 等 11 个基因，编码尾部蛋白质。

3) 重组基因：行使整合、重组和切割功能。含有 *int*，*xis*，*exo*，*red*，*gam* 基因及 *att* 识别位点。(int 蛋白能使 λDNA 整合到细菌染色体上；xis 蛋白是一种酶，能把原噬菌体从宿主染色体上切割下来；int 和 xis 蛋白都能识别 *att* 位点；*red* 基因可促进重组)。

4) 正调控基因：包括 N,Q 两个基因。N 蛋白是早期正调控因子，能拮抗宿主的转录终止子(ρ)，能促进 O,P,Q，*red*，*xis*，和 *int* 基因的转录。

5) 负调控基因：包括 *cI*，*cII*，*cIII*，*cro* 等重要基因。(*c* Ⅰ 蛋白与 O_1,O_2 结合，终止 λ 噬菌体全部的基因转录，促进进入溶原途径及维持溶原状态。*c* Ⅱ，*c* Ⅲ 维持免疫性，阻止裂解基因 S,R 的转录，促进 *cI* 的转录。*cro*，*red* 能部分抑制 *cI*，N，*red* 和 *xis* 的转录。*tof* 使 *cI* 基因关闭，使 λ 噬菌体由溶原进入裂解循环。

6) DNA 合成基因：包括 O,P 基因。

7) 裂解基因：包括 S,R 基因控制溶菌作用；R 基因编码溶菌酶。

(2) 负责转录调控及识别位点：O_L,O_R：左向、右向操纵子。P_L,P_R：左向、右向启动子。P_{RM}：阻遏物维持启动子。P_{RE}：阻遏物建立启动子。*att*：整合与切离识别位点。t_L：左向终止子。t_{R1},t_{R2}：右向终止子 1、2。P_R：右向晚期启动子。*cosL*，*cos*R：左端、右端黏性末端。

2. λ 噬菌体的转录　λ 噬菌体的转录过程分为三个阶段(图 5-10)：早期转录(early stage)，迟早期转录(delayed early stage)和晚期转录(late stage)。

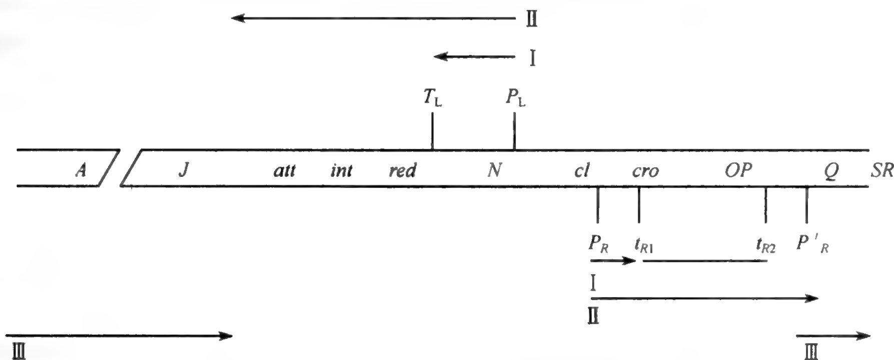


图 5-10　λ 噬菌体 DNA 的转录示意图

Ⅰ. 早期转录本；Ⅱ. 中期转录本；Ⅲ. 晚期转录本．箭头表示转录方向

(1) 早期转录：从 P_L 和 P_R 启动子起始，早期 RNA 的转录分别在 N 基因和 *cro* 基因末端的 t_L 和 t_{R1} 位点终止。转录产生的 N 蛋白可中和宿主 ρ 转录中止因子，使转录越过 t_L，t_{R1} 和 t_{R2} 进入迟早期转录。早期转录产物 cro 蛋白的积累，导致 P_L 和 P_R 启动子关闭。

(2) 后早期转录：主要进行 DNA 复制和重组。当 $\lambda O,P$ 基因的产物合成后，环状 λDNA 开始复制。复制分为两个阶段：早期为 θ 复制，晚期为滚环复制。早期 θ 复制增加了转录和复制的模板数；而滚环复制扩增了大量子代噬菌体 DNA。进入生活周期的循环时，θ 复制停止，滚筒复制则成为主要的方式。

λ噬菌体DNA重组的特征是位点特异性重组,这与λ噬菌体DNA的整合(int)和切离(xis)相关。

(3) 晚期转录:晚期基因区包括与噬菌体头、尾组装和裂解有关的许多基因。Q蛋白的合成会导致晚期转录的开始,P'_R启动子的启动将转录与噬菌体的头、尾组装和裂解有关的基因。然后包装、成熟。

λDNA转录的特点是在生命周期的早期阶段进行的,呈左向、右向双向转录,分别转译、合成相应的酶和蛋白质,基因调控机制也基本明了,其转录调控研究更为彻底。其功能性基因大多聚在一起,编码调控蛋白质的基因与被调控靶位点极其靠近,终止位点与起始位点相重叠。50%或60%的基因对噬菌体的生长和噬斑形成是必需的,但在$J \sim N$之间1/3的基因不是必需的,利用后者的这一特性可构建成不同的克隆载体。

3. 溶原、裂解周期的基因调控　λ噬菌体感染大肠杆菌宿主后,要选择溶原或裂解发育途径。溶原状态需要λDNA整合进宿主的细胞DNA中。由溶原状态进入裂解生长需要λDNA从宿主DNA上切离下来。整合和切离都是在细菌和噬菌体DNA的特定位点经过特异性重组实现的。

λ噬菌体的溶原和裂解循环主要是后早期基因(即调节基因)编码功能的表现,只有它们的功能表达后,使周期循环发生分歧,才显现出溶原或裂解途径。*c*Ⅱ基因是影响溶原和裂解的关键,*c*Ⅱ活化使阻遏蛋白启动子合成有效,则阻遏蛋白结合操纵基因;若cⅡ不活化,阻遏蛋白不能形成,Cro则结合操纵基因。Cro蛋白和阻遏蛋白相互拮抗,Cro蛋白合成可间接阻止*c*Ⅰ转录,而不进入溶原。阻遏蛋白的合成可阻止早期基因的转录,关闭Cro蛋白的合成,使进入裂解周期。λ噬菌体DNA是否进入溶原或者裂解途径,取决于阻遏蛋白和Cro蛋白两者作用的效力与时间。

二、λ噬菌体载体

(一) 基本特性

根据λ噬菌体的基因组特点,利用碱基突变的方法,把λ噬菌体基因组中$J \sim N$之间(约占全基因组1/3长度,是裂解生长非必须区)的DNA序列,用内切酶修饰切点从而将野生型λ噬菌体改造成λ噬菌体突变株(Murray, Rambach, Thomas等,1974)。后又改造成两种λ噬菌体克隆载体:插入型载体(insersion vector)和替换型载体(substitution or replace vector)(Williams和Blattner,1980)。λ噬菌体是最早使用的载体,以溶菌状态生长。其突出优点是:①包装容量大(可达20kb)。②体外包装反应效力高(比质粒载体克隆效率高100倍),故适用于构建cDNA文库和基因文库。

构建λ噬菌体载体的基本条件,是确保其具有感染性和增殖性。要求λ噬菌体基因组的长度适宜,不能过长、过短,否则不能有效包入外壳形成有感染性的噬菌体,一般为基因组的75%~105%之间(即38~52kb)。这对于插入型载体和替换型载体各有特定的要求。λ噬菌体载体的主要优点是体外包装反应非常有效,它比质粒载体克隆效力可高100倍。

(二) 常用的λ噬菌体载体

λ噬菌体载体是载体系列中最为庞大和应用很广的克隆载体。常用的有:λgt、Charon、λEMBL 和 λZAP。按其特点和用途分为两类:插入型载体和替换型载体。

1. λ插入型载体　含一个或多个单一酶切位点,只容许小于 10kb 的 DNA 插入,用插入灭活等检测重组体。根据插入失活的特异性又分为:①功能失活:当 *c* Ⅰ区编码的阻抑物活性被破坏,不能溶原。如 λZAP 载体。②β-半乳糖苷酶失活——如 pBluescript M13。

具代表性的插入型载体有:λgt10(Huynh 等,1985)(图 5-11 上)、λgt11(Young 和 Davis,1983)(图 5-11 下),及其具有体内删除作用的λ噬菌体载体——λZAP。

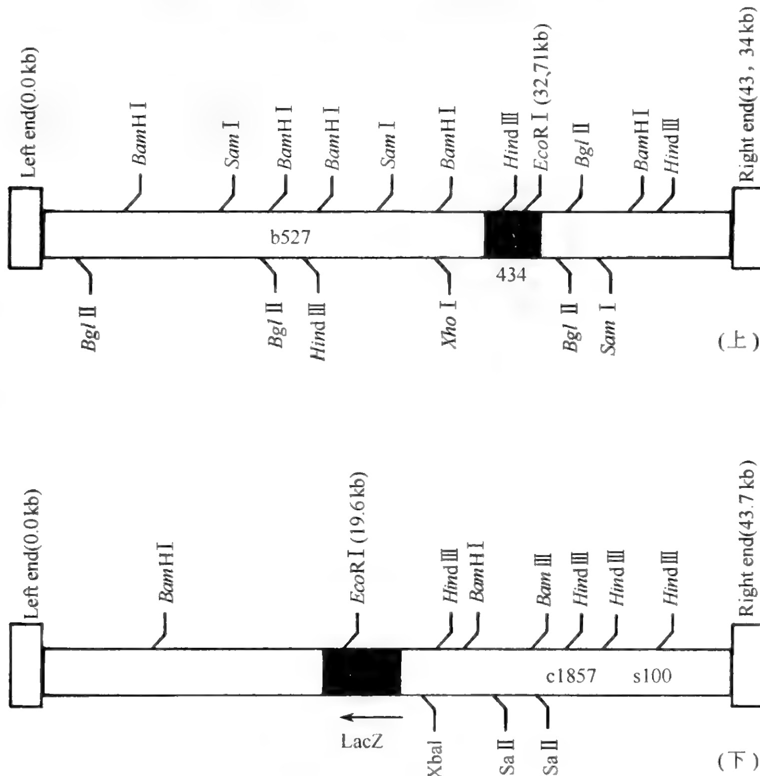


图 5-11　λgt10(上)、λgt11(下)形体图

(1) λgt10:该 DNA 长 43.3kb,能克隆小于 6kb 的 *Eco*R Ⅰ片段。它的 *Eco*R Ⅰ单一切点位于 *c* Ⅰ基因内,该位点插入外源片段会使 *c* Ⅰ失活,重组噬菌体则形成清亮的噬斑;而未插入者则形成浑浊的噬菌斑。

λgt10 的宿主菌是:C600(扩增时用);C600 hf1A(筛选重组体用)。

(2) λgt11:长度为 43.7kb,可克隆小于 7.2kb 的 *Eco*RI 片段。它含有 β-半乳糖苷酶基因 N 端片段,惟一的 *Eco*R Ⅰ切点在 *lacZ*′基因的羧基端。在 gal^+ 大肠杆菌宿主中,重组噬菌体在 X-gal 平板上形成无色的噬菌斑,非重组噬菌体则形成蓝色噬菌斑。λgt11 用于构建 cDNA 文库,若插入方向正确、读码框架正确则可表达融合蛋白(可具有抗原性)。

λgt11 的宿主菌为:Y1090hsdR.

(3) λZAPⅡ载体与 pBluescript M13-噬菌粒:

1) λZAPⅡ载体基因组的基本构成(Short, 1988. 图 5-12):λZAPⅡ载体基因组为 40.82kb,由左至右依次为:*cos* 末端,*A-J* 区,pBluescriptSK 噬菌粒(内含 *lacZ'*基因)和 *MCS* 序列,*c*1857 及 *att*～*cos* 末端;在 pBluescriptSK 噬菌粒内有:① f1(-)起点和起始子、终止子:f1(-)起点区的起始子(Ⅰ信号)可被辅助病毒 M13 的反式作用蛋白识别,启动切割 DNA。②*MCS*:*MCS* 序列置于 *lacZ'*下游,其两端有 T3 和 T7 噬菌体启动子,外源 DNA 插入 *MCS* 可致插入灭活或者合成融合蛋白。③*ColE*1:细菌质粒的复制子。④*Amp*:氨苄西林抗性基因,提供筛选标记。

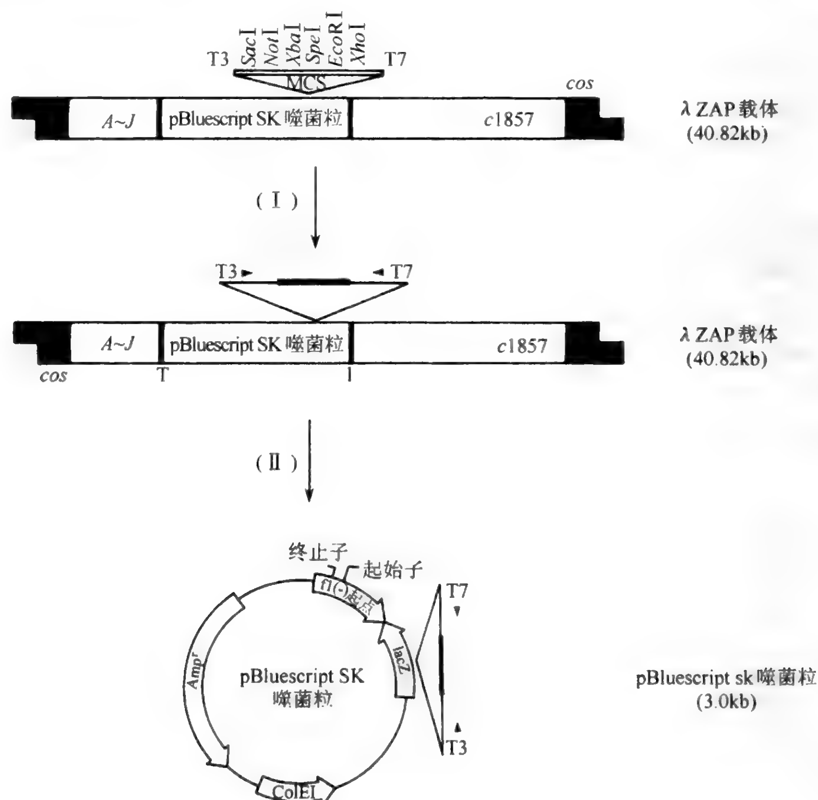


图 5-12 λZAP 载体与 pBluescript M13 的形成过程

2) pBluescript M13 噬菌粒的形成过程(图 5-12):λZAPⅡ载体是一个带有 LacZ-α 互补肽(λZAP,LacZ-α complementation peptide),并可形成 Lac 噬斑的插入型 λ 噬菌体载体,从 *LacZ* 启动子上游 *EcoR* Ⅰ位点插入了一个多克隆位点,在该位点能克隆 10kb 的 cDNA。λZAP 有 6 个限制酶单切点(*Kpn* Ⅰ,*Apa* Ⅰ, *Xbo* Ⅰ, *Hin* Ⅱ,*Acc* Ⅰ,*Sal* Ⅰ,即 K～S 方向),可用于直接克隆 cDNA;在邻近多克隆位点处有 *T*3 和 *T*4 启动子,利用这两个启动子可合成与克隆 cDNA 互补的链特异性 RNA(探针); 另外,该载体内有一个在体内可删除的噬菌粒载体 pBluescript SK(M13-)(Short 等 1988)。f1 序列能在体内将插入的 DNA 从 λ 噬菌

体载体转换到 Bluescript 质粒，经体内删除，包装成单链（+）DNA 噬菌体颗粒——pBluesscriptM13 噬菌粒。最后，用单链（+）DNA 的 pBluesscriptM13 噬菌粒感染 F′大肠杆菌，以 *ColE*1 的 *ori* 起点复制、合成了双股 DNA。

环状双链 DNA 的 pBluescript M13（-）用辅助病毒转染生成 ssDNA，利用此特性能容易地把插入的外源 DNA 测序、位点定向诱变、非定向删除等。λZAP/R 和 λZAP/L 的其他特性完全一样，只是多克隆酶切位点的方向相反，以备选择。其受菌体是商品化的 XL1-BLue MRF 菌株。

3）λZAP 载体的用途：①克隆 cDNA，DNA 长度约 10kb。②有体内删除特性，可直接完成目的 DNA 的亚克隆。③具有多克隆位点和 *lacZ* 插入灭活特性。④可表达融合蛋白，用其抗体筛选。⑤pBluescript M13 能转录 mRNA，可制备外源 DNA 的 mRNA 转录本，而制备 RNA 探针。λZAP 是新一代载体，它增加了克隆位点和表达 cDNA 的范围，因而用途广，有可能替代 λgt10 和 λgt11。

4）应用于：①DNA、RNA 测序，RNA 探针制备。②作单向缺失、定点诱变、表达融合蛋白及制备 RNA 转录物。

2. λ 替换型载体　也称为取代型载体，在其中央有多个限制酶切点，可被外源 DNA 所取代。能克隆长达 9～22kb DNA 片段，多用于构建真核细胞基因组文库，目前常用的该类载体有：Charon 40 和 EMBL4。

（1）Charon 40：Charon 系列最早由 Blattner 等人构建，Charon 40 是近年构建的 λ 置换型载体(图 5-13)。Charon 40 载体全长 41.9kb，左臂 19.2kb、右臂 9.6kb。在中央可置换片段的两侧各有一对方向互为相反的 16 个内切酶的多克隆位点，便于克隆操作。可容纳 9.4～24.2kb 的外源 DNA 片段。在它可置换的片段是由多节段区（polystuffer——由 80 个头尾相接的 DNA 短片段组成），两短片段之间的连接可被 *Nae* Ⅰ识别而切断成小碎片。要清除小碎片，确保大部分噬菌体获得含外源片段的两臂。该重组子是 *gam*$^{+}$，可在 recA^{-}大肠杆菌中增殖。宿主菌为 LE392。

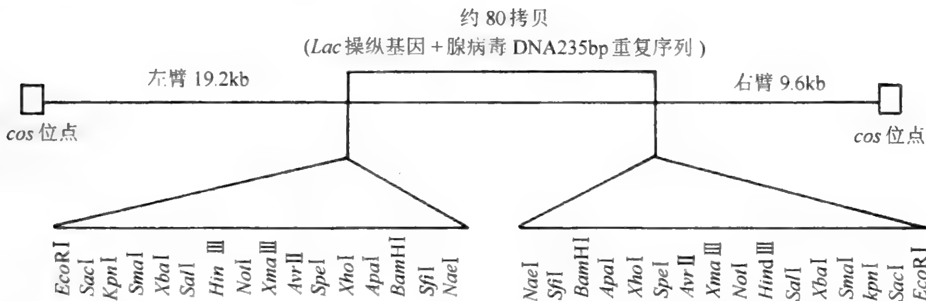


图 5-13　Charon40 载体

（2）EMBL4：EMBL4 是含多克隆位点的 λ 噬菌体置换载体(图 5-14)。由左臂(20kb)、右臂(9kb)、中央区(14kb)三部分组成。在可置换的中间片段，两端有两个切点相同、方向相反的 *Sal* Ⅰ、*Bam*HⅠ和 *Eco*RⅠ多克隆位点，适于克隆 *Sau*3A 或 *Mbo* Ⅰ部分消化的 DNA 片段。被置换的 DNA 序列能抑制 λ 噬菌体的溶原性生长，一旦此片取被外源 DNA 替代，

重组子呈 *Spi* 基因型，能在 P2 溶原化的大肠杆菌菌株中生长良好，产生噬菌斑，利用此 Spi 优势筛选可获得重组噬菌体。

用 EMBL 系列载体克隆重组子，为提高其转染率，转染前需进行体外包装。包装的噬菌体颗粒比未包装的转染率高 100～1000 倍（10^6 噬斑数/每 mg 重组 DNA）。

EMBL4 的受体菌：LE392。

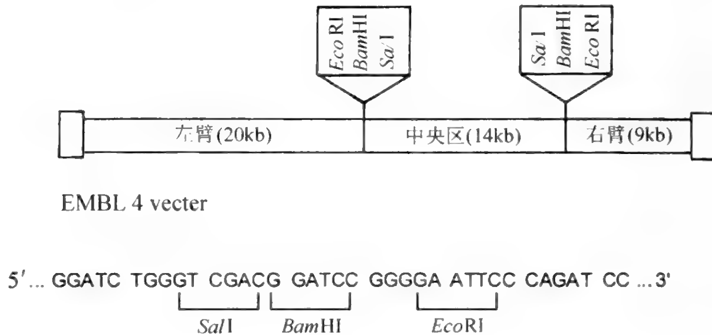


图 5-14　EMBL4 载体

（3）λGEM11：来源于 λEMBL3，中间可取代区的两侧各有 6 个内切酶识别位点（*Sac* Ⅰ、*Xho* Ⅰ、*Bam* HⅠ、*Avr* Ⅱ、*Eco*RⅠ、*Xba* Ⅰ），两侧还有 *Sfi* Ⅰ切点（识别顺序不对称）和方向相对的 SP6、T7 RNA 的启动子。其切割、沉淀、连接、克隆过程与 λEMBL3 相同。

因其插入容量大（24kb），尤适宜建基因组文库。用 Spi^- 筛选，非重组背景稍高。

上述替代型 λ 噬菌体载体，因可容纳 20kb 的 DNA，故较适用于建立基因组文库。应根据实验条件选择合适的 λ 噬菌体克隆载体。

第五节　黏　　粒

一、基 本 特 征

黏粒，也称柯斯质粒（cosmid，来自 cos site-carring plasmid），最初由 Collins 和 Hohn，1978 年以质粒和 λ 噬菌体 *cos* 黏末端序列改建而成。20 世纪 80 年代后有许多改进，扩大了应用范围。

黏粒的基本特性：①质粒的特性：有复制起始点，限制酶切点 1～2 个，耐药性标记；②λ 噬菌体特性：含 *cos* 末瑞序列，基因组长 4～6kb，29～45kb 的目的 DNA 可置于两 *cos* 之间。③克隆容量大（31～45kb）。④能与同源序列的质粒进行重组形成共合体，利用这一特性可把两个筛选标记不同的共合体分子选择出来。把柯斯质粒载体 c2XB 为其中之一，如图 5-15。

黏粒的优势：既具有细菌质粒自主复制的特点，又能包容大片段 DNA（小于 45kb），经体外包装，感染宿主菌，再复制出大量 DNA 或表达产物，易于鉴定。

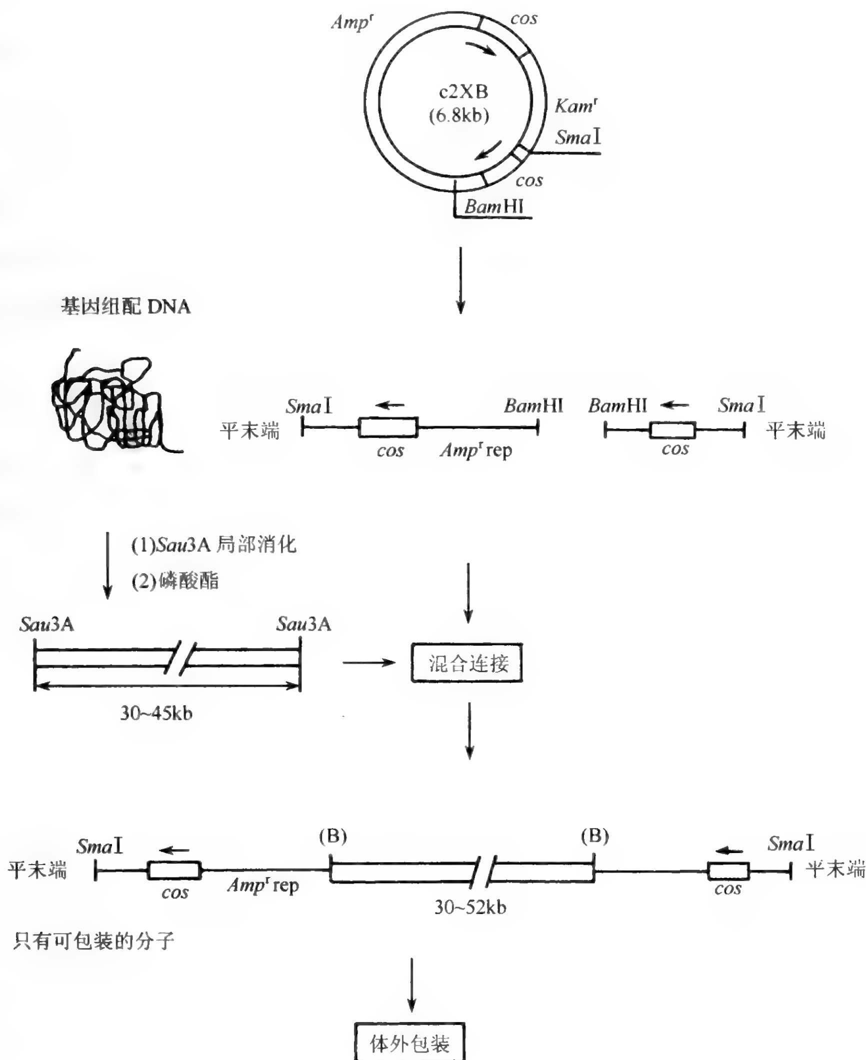


图 5-15 Bafes-Swift 柯斯克隆步骤

二、技 术 改 进

原始的黏粒克隆方法,繁琐、阳性重组率低、易丢失。为克隆真核细胞 DNA,后来的技术改进主要是:

1. 修饰黏粒载体 修饰黏粒载体包括:①使它含有两个 *cos* 位点,便于克隆。如 c2RB 载体。②加入原核细胞的复制子,易在细菌中复制。③加入特异性启动子:*T3*、*T7*、*SP6*,利于双向转录。如 pWE15。④含真核细胞复制起始位点(SV_{40}复制位点),如 pMCS。⑤加入选择标志(*neo* 抗性),利于真核细胞中的转染、筛选。

2. 筛选技术 筛选技术主要为:①菌落原位杂交(大量筛选)。②蛋白质表达(抗原性、

生物活性检测)。③药物抗性筛选。

3. 柯斯克隆方法——Bates-Swift 法(1983)　应用黏粒载体在大肠杆菌中克隆真核细胞基因组大片段 DNA,称为柯斯克隆。

以 Bates-Swift 柯斯克隆法为例(P. F. Bates 和 R. A. Swift,1983),其基本步骤如图 5-15。

(1) 黏粒克隆的基本步骤(以 c2XB 为例,图 5-15):DNA 用 *Bam*H Ⅰ 和 *Sma* Ⅰ 消化后,与经 *Sau*3A 局部消化和碱性磷酸酶处理的真核 DNA 片段重组,此重组分子即作为被包装的底物。包装的噬菌颗粒导入大肠杆菌 $recA^-$ 细胞中,以质粒形式复制,并提供氨苄西林抗性筛选标记。

(2) c2XB 黏粒的主要特性:c2XB 黏粒的特性主要是:①6.8kb 大小,具有 *Bam*H Ⅰ 单一位点,两个 *cos* 位点且被 *Swa* Ⅰ(平末端酶)隔开。②有两个抗性基因(Amp^r、Kam^r; Kam^r 在两个 *cos* 位点之间。③用*Bam*H Ⅰ,*Sma* Ⅰ 两种内切酶双酶消化所得片段与 *Sau*3A 酶切(与 *Bam*H Ⅰ 是同尾酶)的真核 DNA 连接,获得中间为真核 DNA 片段、两端分别为 c2XB 的重组体分子。④最后在体外包装,感染。

(3) 应用:①克隆容量大(47kb)。②适用于建立真核基因文库或者分离大型基因(大于 10.0kb)。

第六节　M13 噬菌体载体

一、野生型 M13 噬菌体的生物学特性

野生型 M13 是一种丝状的大肠杆菌噬菌体,基因组为(+)单链闭环 DNA 分子,6407 bp。它只感染具有 F^+ 性菌毛的大肠杆菌,在宿主细胞内形成双链、环状的复制型 DNA(RF DNA,replication form)。RF DNA 分子以滚环复制累积拷贝。它感染宿主细菌而不裂解宿主菌,但可释放到培养基中,可从培养基中收获该噬菌体。在平板固体培养基上因使细菌生长减慢,形成浑浊的噬菌斑。

M13 噬菌体基因组中,在Ⅱ～Ⅳ基因之间有 50bp 的间隔顺序(IS,intergenic sequence),它不编码蛋白,但与 DNA 复制、转录和包装有关,外源 DNA 插入该区不影响其复制,其他区的基因为复制所必需(图 5-16)。

二、野生型 M13 噬菌体的改造

把野生型的 M13 噬菌体(图 5-16)改造成为 M13mp 克隆载体,主要集中在 IS 区:

1. 插入选择标记　将大肠杆菌 *lac* 操纵子和编码 *lacZ′* 区(表达 β-半乳糖苷酶氨基末端 146 个氨基酸,即 α-肽)能与受体细胞(缺失 *lacZ′* 基因的)F 因子附加体互补,形成有活性的 β-半乳糖苷酶,可借蓝白斑筛选。

2. MCS 插入 α-肽 N 端的编码部分　如果在 MCS 插入 α-肽 N 端的编码部分可得到不同的 M13mp 系列。如:M13mp8/9、M13mp10/11、M13mp18/19(图 5-17)。

三、M13 载体系列的优缺点

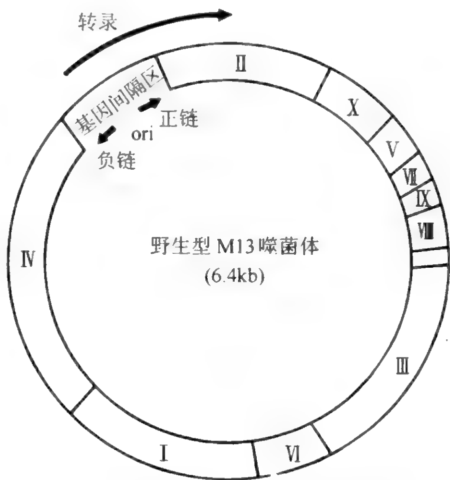


图 5-16　野生型 M13 噬菌体遗传图谱

1. 主要优点　①使克隆操作简便。因有一段密集的序列，插入到修饰改变的 β-半乳糖苷酶基因片段中(*Hin*d Ⅲ片段)。②由于 M13mp 系列是成对地构建的，可有效克隆 ds DNA 中每一条链。③M13mp载体独一无二的特点是产生 ss DNA，这在 DNA 测序、制备 ssDNA 标记探针中得以广泛应用。目前以 M13mp18/19 最常用。

2. 缺点　克隆能力受限，克隆的 DNA 片段不长于 1500bp。

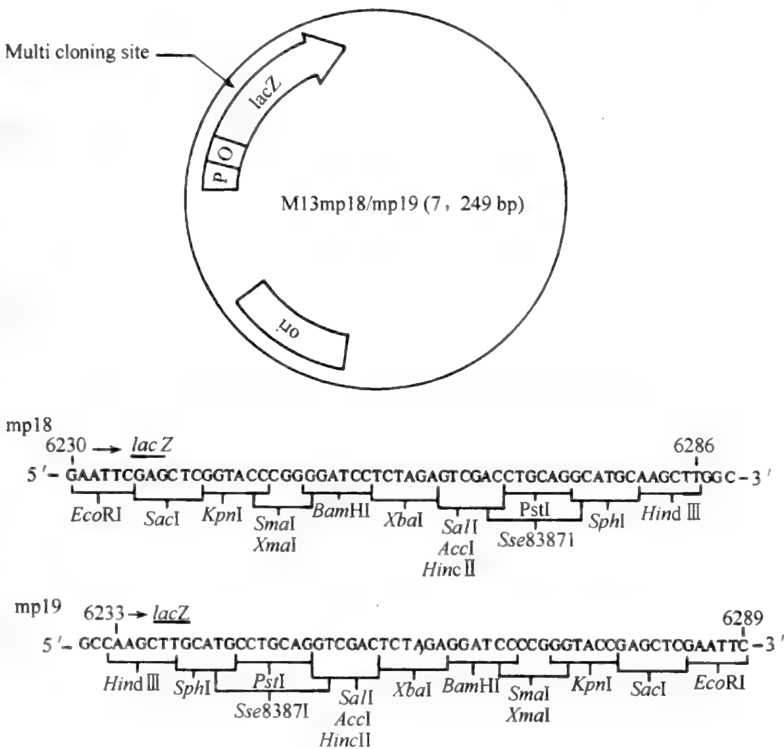


图 5-17　M13mp18/19 限制图谱和克隆位点序列

四、M13mp18/19 的限制图谱和克隆位点序列

1. M13mp 18/19 两者是一对常用的 M13 单链噬菌体载体，基因组为 7.25kb。不同的

是两者 *lacZ′* 区 *MCS* 顺序相反，便于两个方向插入，获得互为相反的两条链。

2. 用 lacZ′α-肽互补选择重组体，即蓝白斑筛选。

3. 有一段 *MCS* 序列，供选择 RE 位点。

五、主 要 应 用

主要应用：①DNA 测序，如 M13 双脱氧酶促法；②制备单链 DNA；③制备单链 DNA 探针。

六、受体细菌株

受体细菌株为：JM103、JM107 和 JM109。

第七节　噬菌粒载体 pUC118/119

一、基 本 特 性

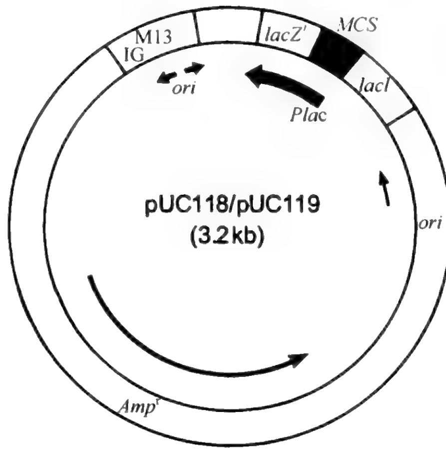
噬菌粒载体是将质粒和单链噬菌体载体的优点结合起来所构成的新型载体体系，称为噬菌粒载体（phagemid 或 vector）。

噬菌粒载体的相对分子质量比 M13 小，约为 3000bp，既有质粒的复制起点，也有噬菌体的复制起点。它在大肠杆菌内可按正常双链 DNA 分子的复制方式形成双链 DNA，具有常规质粒的特性；当在辅助噬菌体（基因Ⅱ蛋白质）的协助下，它又可同 M13 载体一样按滚环模型复制出单链的 DNA，经包装成噬菌体颗粒后被挤压出宿主细胞。利用后者的这一特性，可直接把克隆的 DNA 片段测序。例如 pUC118/119 是较常用的噬菌粒克隆载体。

二、噬菌体载体 pUC118 和 pUC119 的主要特性

噬菌体载体 pUC118/119（图 5-18），是一对分别由 pUC18 和 pUC19 质粒与野生型 M13 噬菌体在基因间隔区（IG，intergenic region）重组构建的。*IG* 区长度为 476bp，含 M13 的复制起点，它插入到 pUC18 和 pUC19 质粒的 Ndel 位点上。两者的 *MCS* 序列正相反，其他分子结构完全相同。如果某一 DNA 分子插入这一对噬菌体载体多克隆位点的同一限制性内切酶位点上，将形成的两个重组载体，一个会转录克隆基因的正链 DNA，另一个则转录克隆基因的负链。因此，应用 pUC118 和 pUC119 噬菌粒载体，能将克隆的 DNA 两条链都能有效地合成出来，可广泛用于 DNA 测序及体外定点诱变等。

但是，要制备单链 DNA，需要让该噬菌体进行滚环复制，所产生的单链载体 DNA 必须由辅助病毒提供外壳蛋白包装，形成噬菌体颗粒再被挤出宿主细胞，分布在培养基中。从该培养基中可获得大量噬菌体颗粒。



A) pUC118 多克隆位点

ATT ACG AAT TCG AGC TCG GTA CCC GGG GAT CCT CTA GAG TCG ACC TGC AGG CAT GCA AGC TTG GCA

*Eco*RI　*Sac*I　*Kpn*I　*Sma*I *Xma*I　*Bam*HI　*Xba*I　*Sal*I *Acc*I *Hinc*II　*Pst*I　*Sph*I　*Hin*dIII

B) pUC119 多克隆位点

ACG CCA AGC TTG CAT GCC TGC AGG TCG ACT CTA GAG GAT CCC GGG GTA CCG AGC TCG AAT TCA CTG

*Hin*dIII　*Sph*I　*Pst*I　*Sal*I *Acc*I *Hinc*II　*Xba*I　*Bam*HI　*Sma*I *Xma*I　*Kpn*I　*Sac*I　*Eco*RI

图 5-18　pUC118 和 pUC119 噬菌体载体

三、辅助病毒 M13K07

pUC118 和 pUC119 噬菌体感染宿主细胞后，只有进入滚环复制，才能有效地包装、释放。而要进入该途径必须有辅助病毒协助才行。pUC118 和 pUC119 噬菌体的辅助病毒是 M13K07 菌株，其基本构成是含有：①M13mp1 噬菌体的复制起点（已突变），使自身的 DNA 不能有效复制，但可在宿主细胞内合成供包装用的外壳蛋白质。②有 p15A 质粒的复制起点，它控制辅助病毒 DNA 复制。③Tn903 的卡那霉素抗性基因。④M13 的 IG 区，它仅有一个 *Ava* Ⅱ 切点，M13mp1 突变的复制起点就插在此区。

四、噬菌体载体 pUC118/119 的优点

1. 具有 pUC 质粒的特性，可用于常规质粒的克隆。
2. 本身相对分子质量小、共价闭合，可克隆 10kb 外源 DNA 分子，容易分离和操作。
3. 有 amp^r 抗性基因选择标记，便于筛选、初步鉴定。
4. 拷贝数量高（可达 500 个/每个细胞），DNA 制备效率高。

5. 有多克隆位点序列，便于选择外源 DNA 与克隆。

6. 载体有 *lacZ′* 基因编码区，*lacZ′* 基因置于 *lac* 启动子控制之下，可用 Xgal- IPTG 组织化学显色反应筛选重组子。

7. 用辅助病毒超感染形成 DNA，能制备单链模板和 DNA 探针，作定点诱变分析等。

本章内容主要讲述了基因工程技术中常用克隆载体的主要特性与应用，现归纳如表 5-2。

表 5-2 常用的基因工程载体的主要特点

载体	基因组大小/kb	宿主细胞	最大负荷/kb	筛选标记	应用			主要特点
					基因库	DNA序列	表达	
质粒	5±	*E.coli*	＜10	耐药性	+	+	+	插入失活，插入表达，体外体内重组相结合
λ噬菌体	50	*E.coli*	23	琥珀，lac 颜色	+	−	−	
黏粒(cosmid)	2～4	*E.coli*	45	耐药性	+	−	−	体外包装 λ-*cos* 尾巴，体外包装像 λ-噬菌体，体内复制像质粒。
M13	6.4	*E.coli*	1～3	lac 颜色	−	+	−	插入失活
pYAC	11.4	细菌，酵母	400	耐药性，基因突变抑制	+	−	−	穿梭质粒，容量大
SV_{40}载体	5.2	猴肾细胞	2.5	敏感细胞	−	−	+	早晚期区域取代，增强子 LTR，cDNA 整合，原病毒
RNA 肿瘤病毒(卸甲载体)	8～10	鼠细胞	7	敏感细胞	−	−	+	
乳头瘤病毒(穿梭载体)	8±	真核：鼠细胞，原核：*E.coli*	?	敏感细胞，耐药性	−	−	+	像质粒一样的附加体
痘病毒	200	敏感细胞	25	蚀斑筛选	−	−	−	宿主范围广，构建多价疫苗
昆虫杆状病毒(转移载体)	80～200	昆虫细胞	100	0	−	−	+	强启动子，体外体内重组相结合

（曹英林　朱法良）

第六章　基因克隆

Chapter 6. Gene Cloning

基因克隆(gene clonging)是通过分子生物学技术获得一段基因或DNA片段的大量拷贝,从而对该基因进行研究或应用的过程和技术,又称为DNA克隆。基因克隆是研究基因组成、结构、功能与调控,以及从分子水平研究各种生命现象,包括生长发育、遗传、进化及疾病等,从而成为改造细胞或个体遗传性状、造福于人类的基础。基因克隆中所扩增的基因称为目的基因。由于基因克隆时必须将目的基因重组至载体中,并导入受体细胞复制,因而目的基因又称为外源基因(foreign DNA)。

基因克隆是分子生物学的核心,其关键技术是DNA重组技术。DNA重组(DNA recombination)即用酶学的方法,在体外将不同来源的DNA进行特异性切割,并重新连接,组合成一个新的杂合DNA分子的过程。

第一节　目的基因的获得

基因克隆的第一步即获得目的基因。根据研究对象和目的的不同,目的基因可来自原核细胞或真核细胞。原核细胞基因组比较简单,其基因定位相对容易,一般可直接分离基因组DNA,酶切后,以相应探针调取目的基因。真核细胞基因组相对复杂,多含有庞大数目的基因,其基因定位困难,真核基因的克隆多通过构建基因组文库进行。

根据人们对目的基因序列的认识情况,可采用不同的基因克隆策略。

一、已知基因的获得

已知基因即序列已被克隆、测序,并做到资源共享的基因。随着分子生物学技术的发展和人类基因组计划的实施与完成,人们对基因的认识正在以前所未有的速度发展,已知基因越来越多。

由于已知基因的一级序列清楚,其基因的获得相对容易,通常可采用化学合成法和PCR扩增法获得。

(一) 化学合成法

该法是Khorana于1967年首先发明的一种利用化学反应合成寡核苷酸链的方法。随着化学合成基因技术的日益完善和自动化,合成基因的周期越来越短,合成的基因越来越大。自20世纪80年代以来,DNA化学合成方面取得了突破性的进展。经过几十年的发

展，已从费时、费力，又难掌握的液相合成法发展为自动化的DNA合成仪，现在计算机控制的全自动核酸合成仪已被广泛应用，可按人们设计好的序列一次合成100～200bp长的DNA片段，这使得DNA来源不再受限制，极大地促进了分子生物学的发展，迄今已有100多种合成基因，如胰岛素、α-干扰素、尿激酶基因等。

目前，全自动合成仪化学合成DNA采用的是效率极高的固相亚磷酸三酯法。

固相亚磷酰胺三酯法是将DNA固定在固相载体上完成DNA的链合成，其合成方向是由合成引物的3′端向5′端合成，相邻核苷酸通过3′→5′磷酸二酯键连接。其基本过程包括：第一步，将预先连接在固相载体上的活性基团被保护的核苷酸与三氯乙酸反应，脱去其5′-羟基的保护基团Dmt，获得游离的5′-羟基；第二步，合成DNA的原料（亚磷酰胺保护的核苷酸单体）与活化剂四氮唑混合，得到反应活性很高的核苷亚磷酸活化中间体（其3′端被活化，5′-羟基仍然被Dmt保护），与溶液中游离的5′-羟基发生缩合反应；第三步是带帽（capping）反应，缩合反应中可能有极少数5′-羟基没有参加反应，带帽反应中使用乙酸酐和1-甲基咪唑使这些核苷5′-羟基反应形成酰，终止其继续发生反应而成为短片段，这种短片段可以在纯化时分离掉；第四步，缩合后在氧化剂碘的作用下，亚磷酰形式转变为更稳定的磷酸三酯。经过以上四个步骤，一个脱氧核苷酸被连接到固相载体的核苷酸上。再以三氯乙酸脱去它的5′-羟基上的保护基团Dmt，重复以上步骤，直到所有要求合成的碱基被接上去。

以上合成的粗产物，使用浓氨水在55℃，8～10h脱去碱基环上的保护基团，再使用各种纯化方式去除合成过程中形成的短片段，得到合乎要求的DNA片段。

DNA化学合成常用来合成PCR引物、DNA探针、测序引物及目的基因合成等，几乎涉及整个生命科学领域。

化学合成法的准确性极高，合成速度较快，是一个十分有用的手段，它不仅可直接合成基因，还能改造基因结构，提高表达效率，化学合成的寡核苷酸DNA片段可作为探针引物和人工接头，在基因工程研究中具有重要的作用。但化学合成价格昂贵，所以目前很少全部用化学方法去合成大片段基因。

（二）PCR扩增

对于已部分了解或完全清楚的基因，可以通过PCR反应直接从染色体DNA或cDNA上高效快速地扩增出目的片段；只要了解基因两侧的DNA顺序，并分别在5′端和3′端设计引物，以含有微量目的基因的样品为模板，利用DNA聚合酶，根据DNA复制的碱基配对原则，可在体外按2^n指数扩增，获得所需基因。PCR的原理及技术参见第九章。另外，最近几年发展起来的反转录聚合酶链反应（reverse transcription-polymerase chain reaction，RT-PCR）已成为克隆真核基因常用的有效方法。该法省去了cDNA文库的构建和筛选等操作，可用于基因转录产物分析、cDNA克隆、cDNA探针制备和RNA高效转录系统的构建。

PCR可以扩增各种材料的DNA，而不必分离目的基因，其成功的关键在于引物的设计。PCR具有简便、快速的特点，目前除用做扩增已知基因外，已成为检测特定基因、进行基因突变分析的重要方法。

然而，由于PCR每轮扩增的产物是下一轮复制的模板，随着快速复制次数的增加，其错误几率也越来越大，实验统计表明，35轮复制后的错误几率约为0.25%，而且由于较小的

DNA片段容易扩增而优先聚集，所以PCR法不适于扩增基因用于研究其精确结构、调节控制位点。

二、未知基因的获得

尽管人类基因组计划已完成，但真核生物众多，真核基因组信息繁杂，尤其是各种功能基因仍不为人所知，因而尚有许多目的基因序列未知。对于未知基因的获得，可以采用下列方法：

（一）构建基因组文库，筛选目的基因

基因组DNA文库（genomic DNA library）理论上包含某一生物基因组DNA的全部序列。

构建基因组文库的基本方法是：从生物组织细胞中提取出全部DNA，用物理方法（超声波、搅拌剪力等）或酶法将DNA降解成预期大小的片段，然后将这些片段与适当的载体（噬菌体、黏粒或YAC、BAC载体）连接，转入受体细菌或细胞，这样每一个细胞接受了含有一个基因组DNA片段与载体连接的重组DNA分子，而且可以复制扩增，许多细胞一起组成一个含有基因组各DNA片段克隆的集合体（图6-1）。

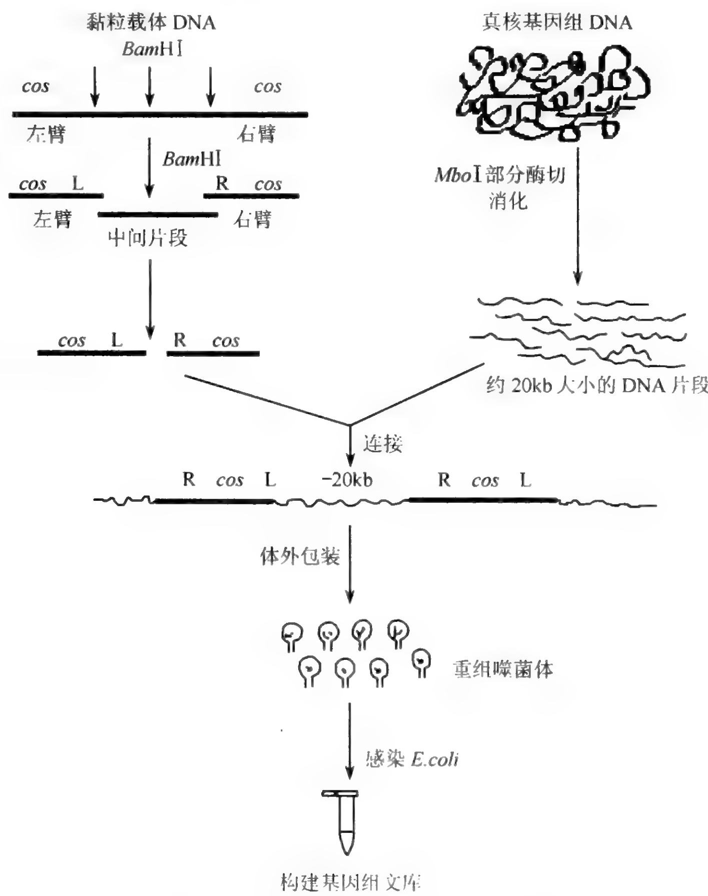


图6-1　基因组文库构建流程示意图

从文库中利用杂交或 PCR 方法能钓取该生物的全部基因或 DNA 序列。

构建基因组文库，再用分子杂交等技术去钓取基因克隆的方法，称为鸟枪法。当生物基因组比较小时，此法较易成功；当生物基因组很大时，构建完整的基因组文库和从庞大的文库中克隆目的基因具有一定难度，因而限制了其应用。

（二）构建 cDNA 文库，筛选目的基因

cDNA 是指以 mRNA 为模板，在反转录酶的作用下形成的互补 DNA（简称 cDNA）。cDNA 文库（cDNA library）即包含着细胞全部 mRNA 信息的 cDNA 克隆的集合，其基本构建过程为：提取组织细胞的全部 mRNA，在体外反转录成 cDNA，与适当的载体连接后转化受体菌，则每个细菌含有一段 cDNA，并能繁殖扩增，这些细菌的集合构成该组织细胞的 cDNA文库。

基因组含有的基因在特定的组织细胞中只有一部分表达，而且处在不同环境条件、不同分化时期的细胞其基因表达的种类和强度也不尽相同，所以 cDNA 文库具有组织细胞特异性。而且 cDNA 文库去除了不编码的非结构基因（如内含子、启动子），比基因组 DNA 文库小得多，能够比较容易地从中筛选出克隆细胞特异表达的基因，是发现新基因和研究基因功能的基础工具。

建立 cDNA 文库后，如何有效地获取或筛选到目的基因，可根据基因而采取不同的方法。若对目的基因序列有所了解，筛选的最直接方法是用核酸探针进行杂交，或设计引物进行 PCR 扩增。若目的基因序列不清楚，而基因产物的部分氨基酸序列已知时，可利用氨基酸序列指导合成核苷酸探针获得目的基因。此外，也可利用差异显示技术筛选组织特异性表达的功能基因。

（三）mRNA 差异显示技术筛选差异表达基因

通过建库的方法筛选目的基因工作量大，成功率低，尤其是随着人类基因组计划的实施与完成，越来越多的基因序列被人们发现和认识，通过建库的方法寻找新基因，几率越来越小。

差异显示是利用一对组织基因表达的差异，筛选出在某种组织中受某因素调控的一个或一组差异表达基因的常用方法。

差异显示技术进展迅速，其中 mRNA 差异显示是筛选差异表达基因的最有效方法之一。mRNA 差异显示又称为差异显示反转录 PCR（differential display RT-PCR，DDRT-PCR），由哈佛大学医学院 Peng Liang 和 Arthar B. Pardee 博士于 1992 年发明建立，至 1996 年发展成熟。

差异显示技术的基本原理和过程是：建立一组具有可比性的细胞（如肿瘤组织和正常组织），抽提其在某一条件下表达的所有 mRNA，构建 cDNA 文库；同时以实验组（肿瘤细胞）和对照组（正常细胞）mRNA 为模板，经 RT-PCR 制备放射性 cDNA 探针，并分别与上述实验组 cDNA 文库进行杂交；然后用变性聚丙烯酰胺测序胶分析其差别，将有差别的基因克隆化，进一步分析其结构与功能（图 6-2）。

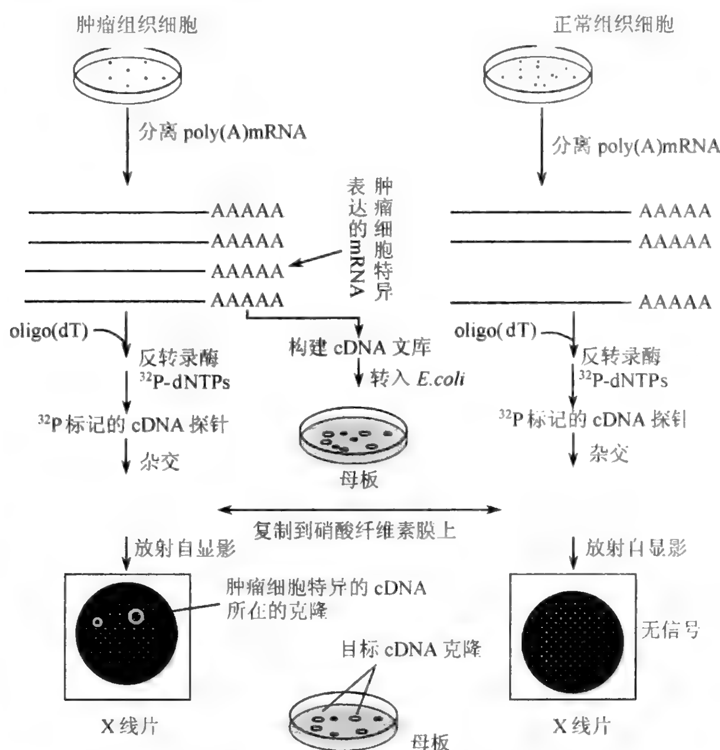


图 6-2 用差异显示杂交筛选肿瘤细胞特异表达的基因

差异显示技术可以筛选同一种细胞在不同条件下表达的基因(包括组织特异性表达,组织发育或细胞周期的某一阶段表达的已知或未知基因),也可以筛选受某一刺激因素或病理因素调控表达的基因。差异显示主要用于肿瘤等多种疾病的分子遗传学研究,可有效筛选获得疾病相关分子,也可用于对不同类型、不同时期细胞的基因或基因表达情况进行比较,以获取对整个细胞生命过程的有用信息。

差异显示的优点在于:①简单,仅需要 PCR 及 DNA 测序胶电泳,可在大多数实验室开展;②灵敏性高,仅需 0.2μg RNA 作为起始材料;③可同时分析多组样品;④重复性好,90%的条带可重现;⑤快速,2 天内可获得 cDNA 谱带,一周内可进行探针的扩增及杂交筛选。由于上述优点,差异显示广泛应用。但差异显示技术也存在不少问题,概括起来有两方面:一是假阳性多(约 50%～70%),二是得到的有差异 cDNA 片段短(约 110～500bp)。

(四) 差异蛋白质谱表达技术筛选功能基因

差异蛋白质谱表达技术是利用双向蛋白电泳对相关细胞中蛋白表达的差异进行直接比较,克隆,从而鉴定相关功能蛋白的方法。由于差异蛋白质谱表达技术直接分析蛋白表达的差异,能够最大限度地减少假阳性率,提高工作效率,较 mRNA 差异显示技术更为简便、有效,是功能基因组学研究中兴起的新技术。

差异蛋白质谱表达技术的基本过程包括:①建立具有可比性的细胞在某一条件下的表达

蛋白质谱,即抽提两种或两种以上的细胞中表达的蛋白,经双向蛋白凝胶电泳,将各种蛋白按相对分子质量和带电荷情况完全分开;②比较两种细胞中蛋白表达的差异,通过图像分析和软件处理,发现表达有差异的蛋白质;③克隆差异表达基因。切下差异表达的相应胶带,进行氨基酸测序,获得新基因的部分序列;据此合成探针,即可从相应文库中获取、克隆全基因。

第二节 基因重组

将目的基因插入载体的过程,即基因重组(gene recombination),其基本过程包括:首先在目的基因和载体两端造成切口,然后依赖双链 DNA 黏末端单链序列的互补结合和 DNA 连接酶将切口补上。这一步的实质就是将两种 DNA 在体外进行酶促连接,最终获得一个重组子。正确的克隆需要有正确纯净的 DNA 片段、相对应的末端浓度和适宜的连接物(载体)分子构型。一般连接反应中外源 DNA 与载体的比例采用 3:1,4:1 或 5:1。

不同来源、性质的外源片段采用的连接方法不同。常采用的连接方法有:黏末端连接法、平端连接法、人工接头法和同聚物接尾法。

一、黏末端连接

1. 全同源黏末端连接　如果目的序列两端有与载体上相同的限制性核酸内切酶位点,则同一限制性核酸内切酶酶切后在目的基因与载体的两端产生完全同源的黏末端,在降低温度退火时,能重新互补结合,在 DNA 连接酶催化下,目的序列与载体 DNA 连接,实现基因重组,称为全同源黏末端连接(图 6-3)。

不同的限制性内切酶酶切,如果产生的 DNA 黏末端相同,也同样可用此法连接,如识别 6bp 序列的 *Bam*H Ⅰ 和识别 4bp 序列的 *Mbo* Ⅰ 切割 DNA 后,都产生 5′突出黏性末端 GATC,可以互补结合连接。

全同源黏末端连接是最方便的克隆方法,其连接效率高,一般可采用低浓度连接酶在低温(低于 16℃)、高浓度 ATP 的情况下进行连接。

全同源黏末端连接的缺点是:容易出现载体自身环化、双向插入(即目的基因以两种方向插入载体)和多拷贝插入的现象,从而在转化后出现高背景,使得筛选更为困难。双向插入尽管对基因克隆无影响,却会影响基因的表达。采用碱性磷酸酶事先对载体 DNA 进行去磷酸化处理,可有效避免载体的自身环化现象;建立连接体系时,若将目的基因和载体 DNA 的摩尔数控制在 2～3:1,则可有效减少多拷贝插入的问题。

2. 定向克隆　使目的基因按一定方向插入载体的克隆方案称为定向克隆。最常用的定向克隆方案是用两种限制性核酸内切酶切割载体和目的基因,从而在载体和目的基因两端产生非同源互补的两个黏末端(黏-黏连接)。定向克隆也可以通过在一端造成平端,另一端产生同源黏末端实现黏-平连接。

定向克隆有效地限制了载体的自身环化,并且实现了目的基因的定向插入,在基因克隆,尤其是表达某一基因时,极为有用。定向克隆的连接体系与相应的黏末端连接或平端连接相同。

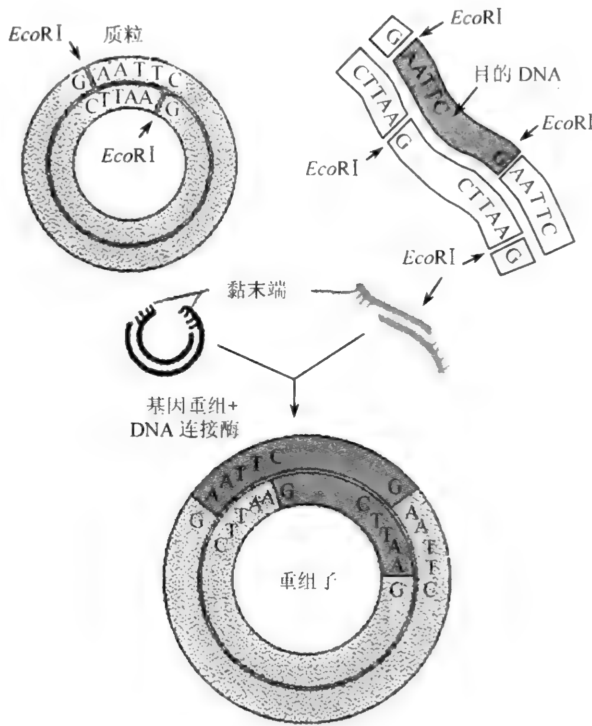


图 6-3 全同源黏末端连接

二、平端连接

T4DNA 连接酶能催化限制性内切酶切割产生的 DNA 平末端进行连接。如果目的序列和载体上没有相同的限制性内切酶位点可供使用,用不同的限制性内切酶切割后的黏性末端不能互补结合,则可用适当的酶将 DNA 突出的末端削平或补齐成平末端,再用 T4DNA 连接酶连接。

平端连接比黏末端连接的效率低很多,但具有普适性,可适用于任何 DNA 片段间的连接,是非常有用的基因克隆策略。另外,平端连接还常用于将人工合成的人工接头连接到 DNA 末端,为进一步的克隆奠定基础。

平端连接的主要缺点是连接效率低和存在多拷贝插入,为此,在平端连接时,需采用高浓度的 DNA 和连接酶,同时,在连接体系中加入低浓度的聚乙二醇类物质,也可有效提高连接效率。多拷贝插入的缺点,可通过控制插入片段与载体的摩尔比解决。

三、同聚物加尾法

如果连接的两个DNA片段没有能互补的黏末端,可通过同聚物加尾法在其末端引入

互补黏末端，即用末端核苷酸转移酶催化脱氧单核苷酸添加至 DNA 的 3′末端，例如一股 DNA 的 3′端加上 polyG，另一股 DNA 的 3′端加上 polyC（图 6-4），这样就可人工在 DNA 两端做出能互补的核苷酸多聚物黏性末端，退火后结合连接，这样的方法称为同聚物加尾法。

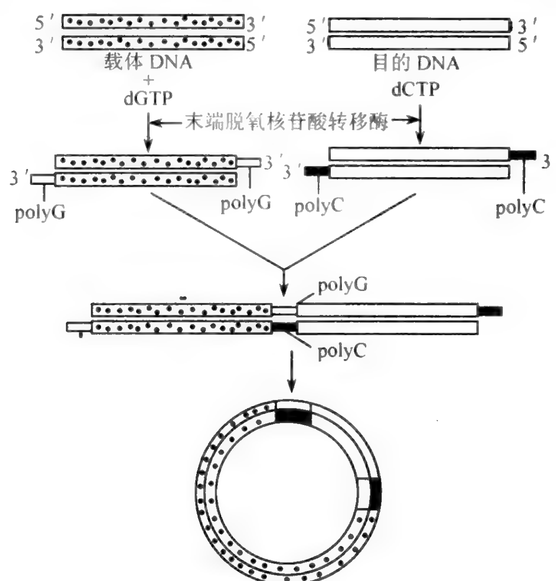


图 6-4 同聚物加尾法连接原理示意图

四、人工接头连接

对平末端的 DNA 或没有互补黏末端的 DNA，除同聚物加尾法外，也可先连上人工设计合成的脱氧寡核苷酸双链接头，使 DNA 末端产生新的限制性内切酶位点，经内切酶切割后，即可按黏性末端相连。人工接头是人工合成的短 DNA 序列，可用于 DNA 的连接，人工接头有两种：连接子和普适子。

1. 连接子（linker） 连接子是一段人工合成的短 DNA 序列，其中含有一或多个限制性核酸内切酶位点，含多个限制性核酸内切酶位点的连接子称为多聚连接子（polylinker）。连接子两端均为平末端，可与目的 DNA 通过平端连接，经合适的酶切后可在目的基因与载体间人为引入黏末端，提高连接效率。

连接子适用于对 DNA 平末端的改造和连接。除此之外，在构建基因组文库时，需要对内切酶消化后的 DNA 片段进行 PCR 克隆、扩增，这时可首先将带有黏末端的 DNA 片段补平，再与连接子连接，这样就在所有 DNA 片段的末端引入了一段共有序列，其中包含了一个相同的内切酶位点。对这样的 DNA 群，只用一种引物（与相应连接子互补）即可对所有 DNA 片段进行扩增，扩增产物经相应的酶切后，可以非常方便的进行克隆。

2. 普适子（adaptor） 普适子是一段人工合成的一端或两端为黏末端的 DNA 序列，其中包含一至多个限制性核酸内切酶位点。普适子克服了连接子仅适用于对平端进行改造的

缺点,可以平端与目的基因连接,也可以黏末端与目的基因连接而改变其酶切位点。使得克隆方案有更多的选择。如：

平末端/*Eco*RⅠ黏末端：

5′ AGCGGCCGCG　　　3′
3′ TCGCCGGCGCTTAA　5′

*Bam*HⅠ/*Pas*Ⅰ黏末端：

5′ GATCCGGCAACGAAGGTACCACTGCA　3′
3′ GCCGTTGCTTCCATGGT　　　　　　5′

五、T-A 克 隆

随着 PCR 及其相关技术的不断发展和完善,其应用范围日益广泛,如何方便、快捷和高效地克隆 PCR 产物越来越成为众多研究者关注的焦点。克隆 PCR 产物有若干好处,一是有利于测序,二是方便产物的大量制备,从而满足进一步研究的需求,如探针标记杂交等。

目前,T-A 克隆是用来直接克隆 PCR 产物的方便方法。其基本原理是:由于 *Taq*DNA 聚合酶在扩增目的基因的同时,可以不依赖模板而在产物的两端加上两个游离的"A",因而只要载体切口处人工加上游离的"T",PCR 产物即可与载体连接。T-A 克隆载体已被广泛应用,并已成为商品化的载体。图 6-5 是 pGEM-T 载体的物理图谱。

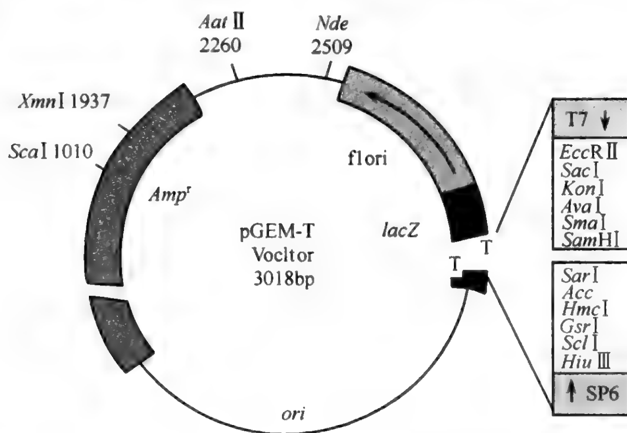


图 6-5　是 pGEM-T 载体的物理图谱

第三节　重组子导入受体菌

目的基因与载体连接后,要导入细胞中才能复制扩增,再经筛选,才能获得重组 DNA 的分子克隆。不同的载体在不同的宿主细胞中复制,导入细胞的方法也不相同。

一、转　　化

转化(transformation)是指将质粒 DNA 或以质粒为载体构建的重组子导入细胞的方法。

早在 1943 年，Avery 等就发现有毒肺炎链球菌的 DNA 与无毒肺炎链球菌共培养后产生有毒性的肺炎链球菌后代的转化现象，证明质粒 DNA 具有进入细胞的能力，但 DNA 自然进入细胞的效率很低。

在分子生物学和基因工程工作中常采取一些方法处理细胞，使之容易接受外界 DNA，处于这种状态的细菌即称为感受态(competent cell)。将感受态细菌与外源 DNA 接触，可大大提高转化效率。最常用的转化方法是 $CaCl_2$ 法，其基本过程包括：首先将大肠杆菌经冰冷 $CaCl_2$ 的处理，成为感受态细菌，然后加入重组质粒并迅速由 4℃ 转入 42℃ 做短时间处理(热休克)，质粒 DNA 就能进入细菌。此外，用高电压脉冲短暂作用于细菌也能显著提高转化效率，即电穿孔(electroporation)转化法。

二、感　　染

噬菌体进入宿主细菌或病毒进入宿主细胞中复制的天然过程就是感染(infection)。

以经人工改造的噬菌体或病毒做载体时，以其 DNA 与目的序列重组后，需经体外包装形成有活力的噬菌体或病毒颗粒，模仿天然过程，借用噬菌体或病毒的外壳蛋白将重组 DNA 注入细菌或细胞，使目的序列得以复制繁殖，这一过程亦称感染。感染的效率很高，但 DNA 包装成噬菌体或病毒的操作较繁琐。

三、转　　染

将噬菌体、病毒或以其为载体的重组 DNA 导入细胞的过程，称为转染(transfection)。

转染不同于感染之处在于，转染时勿需体外包装形成病毒颗粒，而是先经过 $CaCl_2$、电穿孔等处理宿主菌，使之成为感受态，然后，重组的噬菌体或病毒 DNA 也可以质粒 DNA 的方式进入宿主菌，并复制扩增。

最经典的转染方法是 1973 年建立的磷酸钙法，其基本原理是：将 DNA 与磷酸钙混合，制备磷酸钙-DNA 共沉淀物，此时，培养细胞摄取 DNA 的效率会显著提高。此外，用电穿孔法处理培养的哺乳类细胞也能提高细胞摄取 DNA 的能力，但所用外加电场强度、电脉冲时间等条件均与处理细菌不同。脂质体(liposome)转染法用人工脂质膜包裹 DNA，通过胞膜融合将 DNA 导入细胞，方法简单有效，近年来使用日益广泛。

第四节　目的基因序列克隆的筛选与鉴定

目的序列与载体 DNA 正确连接，并重组导入细胞的频率极低。一般一个载体只携带

某一段外源 DNA,一个细胞只接受一个重组 DNA 分子,所以筛选(screening)是基因克隆的重要步骤。筛选目的基因就是挑选出含有目的序列的重组体(recombinant)。

以下就常用技术的基本原理加以介绍。

一、根据重组载体的标志作筛选

克隆载体中都带有用于筛选的标志基因,利用这些标志可以筛选获得阳性重组子。克隆载体携带的最常见标志是耐药性标志,如抗氨苄西林(amp^r)、抗四环素(tet^r)、抗卡那霉素(kan^r)等。若外源基因插入在抗性基因外部,在含有抗生素的培养基中,只有携带相应耐药性基因载体的细胞(阳性重组子)才能生存繁殖,而未能接受载体 DNA的细胞则全部被筛除掉。如果外源目的基因插入在载体的耐药性基因内部,该耐药性基因被插入灭活,则含有目的基因的阳性重组子就会丢掉此种抗性。例如质粒 pBR322 含有 amp^r 和 tet^r 两个抗药基因,若将目的序列插入 tet^r 基因序列中,转化大肠杆菌,让细菌在含氨苄西林或四环素培养基中增殖,凡未接受质粒 DNA 的细胞都不能生长;凡在含氨苄西林和四环素的培养基中都能生长的细菌是含有质粒 pBR322 的,但其 pBR322 未插入目的序列;凡在含氨苄西林的培养基中能生长、而在含四环素的培养基中不能生长的细菌就很可能是含有目的序列的重组质粒(图 6-6)。

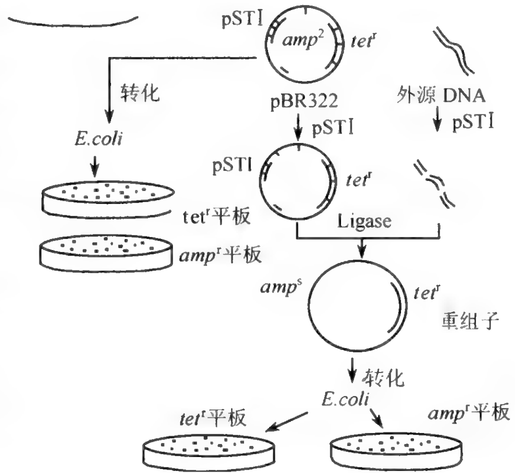


图 6-6 双抗生素对照、插入失活实验

*lac*Z′基因是克隆载体中的另一类重要标志基因,利用 lacZ 设计的蓝白斑筛选法(原理见第五章),近年被广泛应用。例如将目的序列插入质粒 pUC19 的多克隆位点,转化大肠杆菌后,在含氨苄西林、IPTG、X-gal 的培养基中培养,由于 *lac*Z′基因被插入灭活,含有插入目的序列的重组质粒不能产生 β 半乳糖苷酶,呈白色菌落,这样就很容易获得含有目的序列的克隆。

根据重组载体的标志来筛选,可以去除大量的非目的重组体,但只是粗筛,在某些情况会出现假阳性(例如细菌可能发生变异而引起耐药性的改变,却并不代表目的序列的插入),所以需要做进一步的鉴定。

二、DNA 限制性内切酶图谱分析

这是在上述筛选后的进一步分析。目的序列插入载体会使载体 DNA 限制性酶图谱(restriction map)发生变化。利用这种变化可以鉴定插入序列的大小和方向。最常用的方法就是利用重组时的酶切割重组子 DNA,若获得与目的基因一致的片段即证明重组子中含

有插入序列。

三、核酸杂交法

利用标记的核酸做探针与转化细胞的DNA进行分子杂交，可以直接筛选和鉴定含目的基因的阳性克隆。常用的杂交方法包括菌落原位杂交、Southern杂交。其中菌落原位杂交直接将转化后生长的菌落复印到硝酸纤维素膜上，经碱裂解后，将菌落释放的DNA原位吸附在膜上，再与标记的核酸探针温育杂交，核酸探针就结合在含有目的基因的菌落DNA上而不被洗脱，待显色后就可以将含有目的序列的菌落挑选出来。

四、PCR法

PCR技术的出现给克隆的筛选增加了一个新手段。如果已知目的序列的长度和两端的序列，设计合成一对引物，以转化生长的细菌质粒DNA为模板进行扩增，挑选出PCR产物与预期长度相符的克隆，可能就是含有目的序列的重组子。

五、免疫学方法

此法不是直接筛选目的基因，而是通过特定的标记抗体与基因表达产物的反应指示含有目的基因的转化细胞，因而要求目的基因进入受体细胞后能够表达。常用来标记抗体的酶包括过氧化物酶、碱性磷酸酶等。利用酶可催化特定的底物反应而呈现颜色变化，可以指示出含有目的基因的克隆位置。

免疫学方法特异性强、灵敏度高，适用于从大量转化细胞集合体中筛选很少几个含目的基因的细胞克隆。

六、核苷酸序列测定

经上述方法筛选鉴定的克隆，都要用核酸序列测定进行最后鉴定。通过测序可以确证已知基因的克隆准确无误；如果克隆序列未知，通过测序才能确知其结构、推测其功能，用于进一步的研究。因此核酸序列测定是分子克隆中必不可少的鉴定步骤。核酸序列测定的原理和方法详见第十一章。

（马春红）

第七章　原核细胞表达系统

Chapter 7. Prokaryotic Expression System

基因工程的一个重要目的就是获得有功能的蛋白质产品。这一目的的实现,包括目的基因的克隆、复制、转录、翻译、加工、分离纯化等一系列复杂过程。其中,基因的克隆、复制、转录及翻译需要在一个合适的系统中完成,即基因表达系统。根据受体细胞的不同,表达系统分为原核表达系统和真核表达系统,本章主要介绍原核表达系统。

所谓原核细胞表达系统(gene expression in prokaryotic cells)就是将克隆的外源基因导入原核细胞,并使其在原核细胞中以发酵的方式快速、高效地表达、合成基因产物。原核细胞表达系统中用于基因克隆的载体主要是质粒和噬菌体载体(详见前述),外源基因通过转化或转染的方式进入受体细胞,并在其中表达。受体细胞是外源基因最终表达的场所,在原核细胞表达系统中主要指大肠杆菌、链霉菌和枯草杆菌。

本章主要讲述如何将克隆的外源基因导入原核细胞并得到高效表达。

一、外源基因在原核细胞中的表达

外源基因必须在宿主细胞调节元件控制下才能正确表达并获得有效数量的功能蛋白质。原核细胞结构简单、表达调控方式单一,基因工程中很多复杂的基因,尤其是真核生物基因要在其中表达存在一定难度,因而有必要首先讨论一下原核生物的基因结构和表达特点。

(一) 原核生物基因结构和表达特点

1. 基因结构　同真核生物相比,原核细胞没有核膜,基因结构简单。原核基因是连续的,其基因表达以操纵子为单位。操纵子是一些功能相关的结构基因串联在一起,前面结合有调控区控制结构基因的表达,这些结构基因和调控基因共同组成一个完整的转录单位(操纵子),完成一定的功能。

2. 表达特点　原核生物基因的转录和翻译是偶联的、连续进行的,mRNA 在合成过程中可和多个核糖体结合,同时开始翻译多条肽链,即可形成多顺反子 mRNA。原核细胞中缺乏转录后加工系统,不能识别内含子,同时也缺乏真核细胞中的翻译后加工修饰系统。

原核生物中基因的表达,同其他生物一样包括基因的有效转录和 mRNA 的正确翻译。

(1) 基因的有效转录:原核生物中惟一的 RNA 聚合酶在转录过程中识别并与 DNA 启动子结合,启动 DNA 的转录。启动子是一段 DNA 序列,是 RNA 聚合酶识别并结合的部位,对起始 mRNA 的合成必不可少。原核细胞中各种不同启动子都含有两处高度保守区:

一是在转录起始点上游大约 10bp 处，称为 Pribnow box，是一段富含 A、T 碱基组成的序列，称为 -10 区，或 TATA box；另一处是位于转录起始点上游大约 35bp 处，一段由 10 个碱基组成的序列，称为 -35 区。不同启动子的效率不同：RNA 聚合酶很容易识别强启动子，从而产生较多的 mRNA，弱的则相反。-10 区和 -35 区是决定启动子强弱的重要因素，二者间隔的碱基并不十分重要，但二者间的距离（核苷酸数目）却是影响启动子强弱的重要因素。

原核生物中基因转录的终止是以终止子（terminator）来完成的。终止子是位于基因或操纵子 3′ 末端转录终止点之前的一段长约 7～20bp 的特殊序列，由一个富含 A/T 的区域和一个富含 G/C 的回文对称序列构成，回文对称结构的中轴距转录终止点 16～24bp。终止子转录后的 mRNA 有茎环结构，并具有与 A/T 区对应的一串 U（见图 1-10）。由于 A-U 之间氢键结合较弱，因而 RNA/DNA 杂交部分易于分开，从而有利于转录产物从 DNA 模板上释放，并使 RNA 聚合酶脱离 DNA，转录终止。

（2）基因表达的正确翻译：原核生物中转录与翻译是偶联进行的。在 RNA 的合成过程中核糖体即可结合于 mRNA 上，启动翻译过程。原核生物中 mRNA 的核糖体结合位点包括翻译的起始密码子（AUG）及 SD 序列。SD 序列（Shine-Dalgarno 序列）是 1974 年由 Shine 和 Dalgarno 首先发现的，是细菌核糖体有效结合于 mRNA 和翻译起始所必需的核苷酸序列，由 AUG 上游 3～10bp 处的一段 3～9bp 长的富含嘌呤核苷酸的序列组成，该序列刚好与 16S rRNA 5′ 末端的富含嘧啶的序列互补（图 7-1）。



图 7-1 SD 序列的位置

另外，在大肠杆菌中凡发生了无义突变的基因，其编码的蛋白质短肽都被迅速降解，而野生型蛋白质则非常稳定。因而外源基因表达后产生的蛋白质还必须耐受这种降解作用，基因的表达才能最终完成。

（二）外源基因表达的必要条件

综上所述，克隆基因要在原核细胞中获得有效表达，必须具备下列条件：

1. 外源基因的编码区不能被内含子分隔，真核基因的 5′ 非编码区必须被删除。

2. 外源基因必须置于原核细胞的强启动子和 SD 序列等调控元件控制下，启动子必须被 RNA 多聚酶识别。

3. 外源基因在载体中必须有正确的开放阅读框架（open reading frame，ORF）

4. 外源基因 mRNA 必须相对稳定和有效地翻译，形成的蛋白质不易被降解。

二、影响外源基因在原核细胞中表达效率的主要因素

外源基因能否在原核细胞高效表达决定于原核细胞启动子强弱、拷贝数高低及核糖体结合位点等因素。

(一) 启动子

基因高效表达，必须选择强启动子，常用的启动子有：

1. *lac* 启动子　来自大肠杆菌的乳糖操纵子(Jacob & Monod，1961)。乳糖操纵子的组成包括：阻碍蛋白基因(*I*)、启动基因(*P*)、操纵基因(*O*)和结构基因(*Z*、*Y*、*A*，分别编码β-半乳糖苷酶、渗透酶、乙酰化酶)。*P* 是 RNA 聚合酶结合的部位，用于起始转录；*I* 基因编码阻遏蛋白。阻遏蛋白与操纵基因 *O* 结合后，阻止结构基因的转录(图 7-2A)；若诱导物与阻遏蛋白结合，阻遏蛋白不能与操纵基因结合，转录可正常进行，结构基因表达(图 7-2B)。异丙基-β-D-硫代半乳糖苷(isopropyl-β-D-thiogalactoside，IPTG)是 β-半乳糖苷酶的底物类似物，有很强的诱导能力，可与阻遏蛋白结合，促进转录(可提高真核基因表达量 50 倍)。

2. *trp* 启动子　从大肠杆菌色氨酸操纵子中分离，受两种调控：一种相当于 *lac* 启动子的负调控：阻遏基因编码的阻遏蛋白与色氨酸结合后被活化，封闭操纵基因，使转录不能进行；色氨酸缺乏时，阻遏蛋白不能与操纵基因结合，转录被启动。另一种调控通过衰减子完成：色氨酸丰富时，转录进行到启动基因和操纵基因的衰减子停止，而色氨酸缺乏时，转录可以通过衰减子，进行到结构基因。β-吲哚丙烯酸是色氨酸竞争性抑制剂，可以竞争性抑制色氨酸与阻遏蛋白的结合，促进转录。

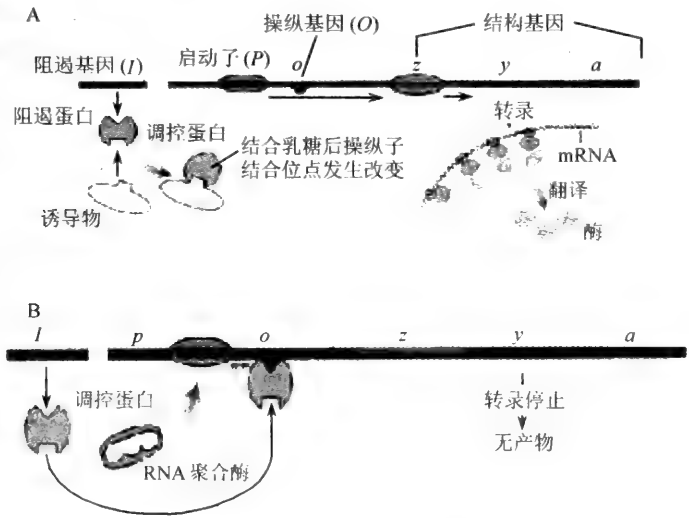


图 7-2　乳糖操纵子的调控模式

3. *Tac* 启动子　是一组 *lac* 和 *trp* 启动子组合的杂合启动子(由 *trp* 启动子的 −35 区和 *lac* 启动子的 −10 区和人工合成的 DNA 及 *lac* 操纵基因组成)。启动能力非常强，受 *lac*

阻遏蛋白的负调节，并被 IPTG 诱导。

4. P_L 和 P_R 启动子　即 λ 噬菌体早期左向和右向启动子。比 *lac* 启动子活性强 8～10 倍。受 λ 噬菌体 *C* Ⅰ 基因负调控，*C* Ⅰ 基因编码温度敏感蛋白，其作用类似于 *lac* 启动子的阻遏蛋白，28～32℃ 时，*C* Ⅰ 基因产生正常的有功能蛋白，抑制转录进行，温度升高至 42℃ 时，*C* Ⅰ 被破坏（可逆过程），操纵基因解除封闭，转录开始。

强启动子可以起始高水平转录，容易导致质粒的不稳定性，在强启动子下游安置一终止子，可以增强质粒的稳定性。

（二）基因剂量

一般来说，细菌内基因拷贝数越多，基因表达量越高，因而可以通过提高基因拷贝数来增加表达量。比如，可以采用温度诱生性载体，这类载体在低温时拷贝数较低，温度升高后拷贝数大量增加，表达量升高。

（三）核糖体结合位点

翻译效率受下列因素影响：①SD 序列与 16S rRNA 3′ 末端的互补程度；②SD 序列的核苷酸组成及其与 AUG 之间的距离；③AUG 两侧位于 −20 至 +13 之间的核苷酸组成。

三、原核细胞表达载体的条件及特点

一段删除内含子和 5′ 非编码区的外源基因经酶切、连接、克隆成功后，要在原核细胞中表达，必须借助于合适的表达载体。表达载体（expression vector）不同于克隆载体，前者是适用于受体细胞中表达外源基因的载体；而后者是适用于受体细胞中复制、扩增外源基因的载体。二者的功能不同，其构成也不同。根据外源基因在原核细胞中表达所需的必要条件和影响表达量的因素，一个成功的原核表达载体必须具备下列元件：

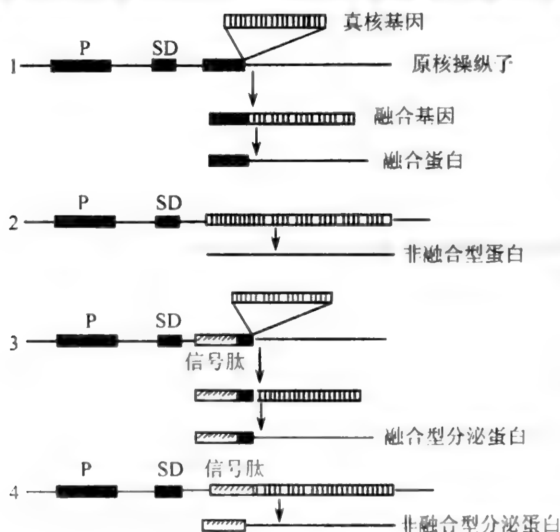


图 7-3　原核基因表达示意图

1. 强启动子　产生大量与克隆基因互补的 mRNA。

2. SD 顺序　提供合适的核糖体结合位点，启动正确、高效的翻译过程。

3. 筛选标志基因和其他一些调控基因　不同载体的调控基因不同，可产生不同的基因产物。

（1）融合型表达载体产生融合蛋白，融合蛋白的一部分（N 端）是由载体编码的原核多肽，而另一部分是一段维持正确 ORF 的外源基因编码的外源蛋白，可通过化学降解法或酶解法切除融合蛋白中的原核多肽而获得外源蛋白。

（2）非融合型表达载体可产生天然完

整的外源蛋白，其中不含任何原核序列。

(3) 分泌型表达载体可将表达的蛋白由细胞质跨膜分泌到胞周间隙，成为分泌型蛋白，分泌型蛋白可以是融合蛋白，也可以是非融合蛋白，三种载体的表达原理见图 7-3。

四、融合型表达载体

融合型表达载体可以将外源基因插入至一段原核序列 3′ 端，最终产生和原核蛋白融合的外源蛋白。

(一) 主要组成元件

融合型表达载体主要由下列元件组成：①原核强启动子和 SD 顺序；②多克隆位点位于一段原核基因的 3′ 端，并使外源基因的 ORF 保持不变。

(二) 构建方法

构建融合型表达载体的关键是将外源基因插入原核序列近 3′端，并保持正确阅读框架，常采用以下方法达到这一目的：

1. 若外源基因和原核基因序列都清楚时，可选择合适酶切位点，使框架正确；

2. 在将外源基因和原核序列连接时，分别将二者末端接上不同长度的、相互匹配的人工寡核苷酸链(dC-dG/dA-dT)形成三种不同阅读框架的重组体，其中之一具有正确阅读框架；

3. 在外源基因和原核序列之间加上合适的人工合成的 DNA 接头，使得阅读框架正确；

4. 同时构建具有三种阅读框架的一系列载体(位相载体)。在这些载体中具有相同的启动子、SD 顺序和一段相同的原核多肽序列，在此序列下游接有一限制性酶切位点，供外源基因插入，此位点距 AUG 之间的距离在一系列载体中各不相同。这样外源基因克隆入酶切位点后，总有一个具备正确阅读框架的表达载体，可以表达外源基因。

(三) 融合型表达载体的特点

目前已构建了大量融合表达载体，很多已商品化，它们具有明显的优越性：

1. 表达效率高　因为外源基因转录和翻译的起始均由正常原核序列控制，故可产生高表达，外源蛋白可通过化学或酶促加工切割得到；

2. 产生的融合蛋白比天然蛋白质更稳定，融合蛋白是避免细菌蛋白酶降解的最好措施；

3. 产生的蛋白质易于鉴定、纯化　融合蛋白相对分子质量大于菌体蛋白，电泳即可鉴定、分离；另外，有些融合蛋白带有特殊功能多肽片段，易于纯化，例如 pGEX 带有谷胱甘肽巯基转移酶(glutanthione *S*-transferase，GST)基因，可利用该酶进行纯化(图 7-4)。

(四) 融合型表达载体举例

pGEX 系统(图 7-4) 包括 3 种载体 pGEX-1T、pGEX-2T、pGEX-3X 和 1 种用于纯化表

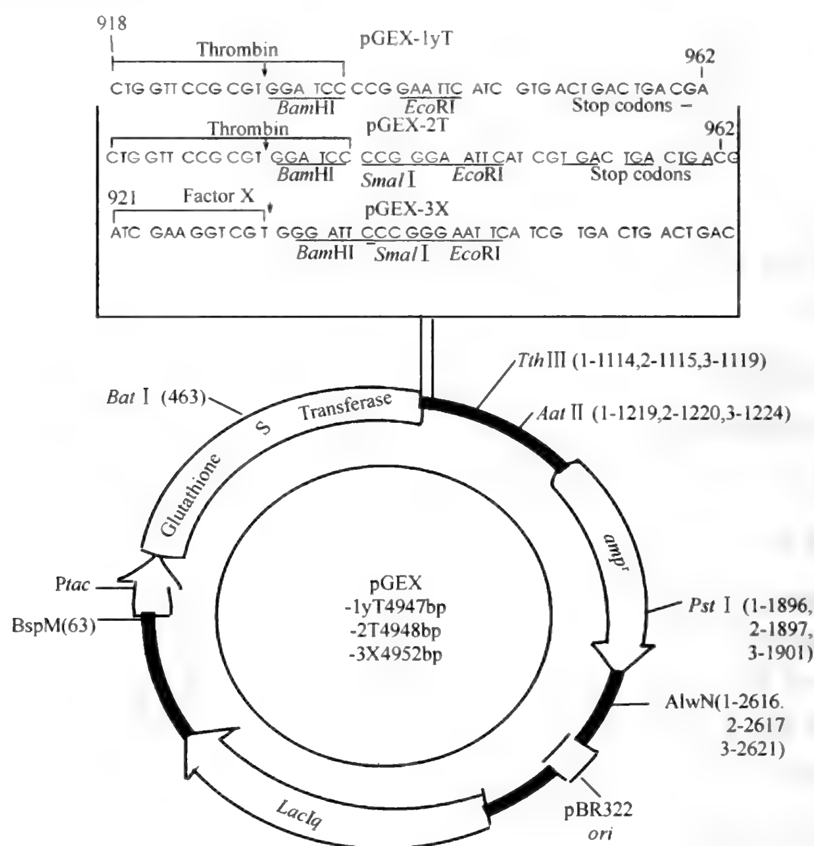


图 7-4 pGEX 载体

达蛋白质的亲和层析介质 Glutathione Sepharose 4B。载体中含有①*tac* 启动子:可诱导高表达;②含有 *lac*I*q* 基因,受体不受限制;③SD 顺序下游为谷胱甘肽巯基转移酶(*GST*)基因,外源基因与 *GST* 基因相连,因而产生的是 GST 与外源基因的融合蛋白,易于纯化(可利用亲和层析介质 Glutathione Sepharose 4B 纯化);④多克隆位点上游含有凝血酶(thrombin)(pGEX-1T、pGEX-2T)或 Xa 因子(pGEX-3X)的特异性位点,因而用凝血酶或 Xa 因子即可获得外源蛋白,对外源蛋白的抗原性和功能影响不大。

五、非融合型表达载体

(一) 主要元件

非融合型表达载体中外源基因从 ATG 开始直接在原核调控元件控制之下,可产生天然完整蛋白。其载体的基本结构包括:原核启动子,SD 序列,ATG,外源基因和终止密码子。

(二) 非融合型表达载体的特点

非融合蛋白在胞内表达时,往往会形成包涵体、包涵体的形成是外源蛋白高效表达时

的普遍现象，其形成与多种因素有关，如：蛋白质的性质、宿主细胞类型、培养温度及营养条件等。尽管包涵体形成有利于产物的纯化，但包涵体的后续处理过程十分复杂，包括：以强变性剂变性溶解包涵体，然后经复性获得有正确折叠的蛋白质。包涵体的复性效率往往较低。

(三) 非融合型表达载体举例

pKK233-3(图 7-5)：可在含 *lac* Ⅰ基因的宿主菌(JM105)中有效表达大量外源蛋白，并可被 IPTG 诱导表达。含强 *tac* 启动子，其下游为一多克隆位点，为稳定载体，在多克隆位点下游还含有一个很强的终止子 *rrnB*，载体的其余部分为 pBR322。

图 7-5　pKK233-3 载体

六、分泌型表达载体

分泌型表达载体将蛋白分泌出胞膜，避免被宿主细胞的蛋白酶降解，是一种十分有用的载体。

(一) 主要元件

1. 启动子和核糖体结合位点；

2. 信号肽序列　为使蛋白质跨过细胞膜分泌至胞周间隙，分泌型表达载体在 SD 顺序下游装有一段信号肽序列。信号肽是一段 15～30 个氨基酸的多肽，具有一个高度疏水核

心，可以引导蛋白质穿过细胞膜，自身被信号肽酶水解，释放出功能蛋白，由于真核生物和原核生物的信号肽十分相似，因而分泌型表达载体中的信号肽序列可取自不同生物细胞。

（二）分泌型表达载体特点

分泌型表达载体产生的蛋白质可分泌至胞周间隙(很少排出细胞，进入培养液)，可避免细胞蛋白酶的降解，并利于纯化。因为除具备信号肽外，蛋白质内部必须有相应的与分泌有关的氨基酸序列，而宿主细胞又具有相应的转运机制，能识别并正确切割信号肽，才能获得蛋白质的分泌表达，所以并非所有蛋白质都可以插入此种载体在原核细胞中得到分泌表达。一般来说，真核生物和原核生物中的分泌蛋白以及一些小分子多肽能在原核细胞中分泌表达，但一些非分泌的天然蛋白则很难分泌表达。

（三）分泌型表达载体举例

1. pINⅢ系统(图 7-6)　包括 3 种载体：pIN-ompA1，pIN-ompA2 和 pIN-ompA3。如图 7-6 所示，三种载体均带有大肠杆菌最强启动子 Ⅰ *pp*(脂蛋白基因)，启动子下游装有 *lacUV*-5 的启动子及其操纵子。为调节基因的表达，在该质粒上克隆有 *lac* Ⅰ 基因，为达到高效表达的目的，质粒中带有人工合成的高效翻译起始顺序(SD 顺序及 ATG)。信号肽序列取自于大肠杆菌分泌蛋白：*opma*(外膜蛋白)基因，信号肽序列下游紧接着是一个人工合成的多克隆位点(*Eco*RⅠ、*Hin*dⅢ和 *Bam*HⅠ)，为使各种目的基因插入多克隆位点后都能保持正确框架，三种载体的多克隆位点片段长度不同，可分别适合于所有三种密码子阅读框架。

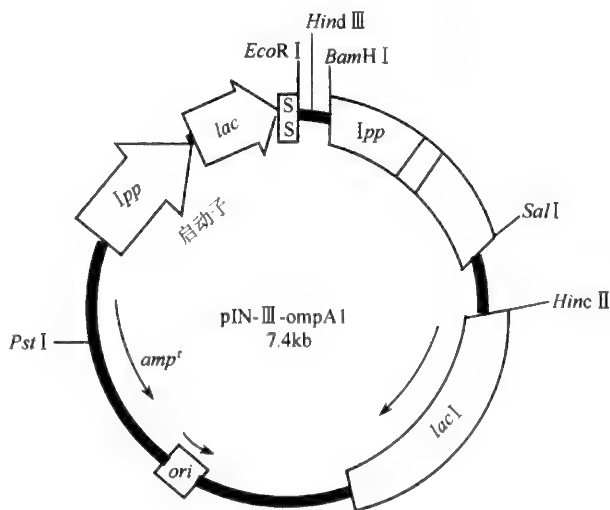


图 7-6　pINⅢ系统

2. pEZZ18　如图 7-7，载体带有 *lac* 启动子，启动子下游装有葡萄球菌 A 蛋白的信号肽序列和两个结构基因 *ZZ*，*ZZ* 基因的下游紧连着一段取自 pUC18 的多克隆位点。因此，pEZZ18 产生与 A 蛋白 ZZ 段融合的蛋白，并可将此融合蛋白分泌至胞周间隙。因为 ZZ 段

是A蛋白与IgG结合的区段，所以该融合蛋白可直接通过亲和层析进一步纯化（以IgG Sepharose 6FF为配体）。另外，由于14kD的ZZ段构型独特，可减少外源蛋白由融合状态恢复为天然状态时受到的影响。

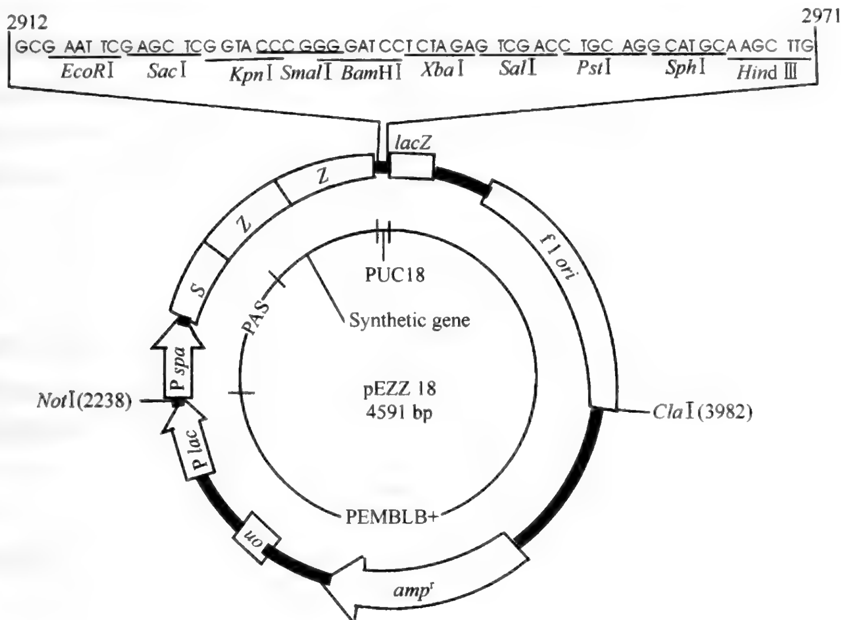


图7-7　pEZZ18

七、表达产物的检测

克隆基因插入适当表达载体，构建表达性重组体，转化受体菌，经诱导表达才能获得表达蛋白，对其表达产物还需进一步检测验证，进而分离纯化。

对于非分泌蛋白来说，所得蛋白质在胞质内，因而检测的第一步就是破碎细胞；对于分泌蛋白来说，蛋白质也仅分泌至胞周间隙，同样要经破碎后才能进行检测，裂解原核细胞的方法主要有超声破碎法、酶解法和化学法。

细胞破碎后，要对蛋白表达产物进行特异性鉴定和生物学活性鉴定。特异性鉴定的方法常用的有免疫沉淀法、免疫荧光素标记抗体法、蛋白质印迹法和ELISA。

（一）免疫荧光素标记抗体法

该法简单快速，适宜于定性检测。

将表达产物的受体细胞固定于干净玻片上，然后与目的蛋白的多抗或单抗37℃充分反应，洗涤后再加荧光标记的第二抗体反应，第二抗体可用荧光标记的葡萄球菌A蛋白或羊抗鼠/兔IgG，充分洗涤后立即用荧光显微镜观察。

（二）免疫沉淀法

可检测并定量分析存在于多种蛋白质复合物中的微量抗原，敏感度极高，可检测到100pg的放射性标记蛋白。

首先以35S-Met和35S-Cys标记受体细胞中的蛋白质；然后将受体细胞裂解，取裂解液（含目的蛋白）与靶蛋白抗体反应，形成抗原-抗体复合物，用偶联的蛋白A和Sepharose或灭活固定的金葡菌（*S. aureus*）吸附抗原-抗体复合物，也可以用抗被检蛋白抗体的抗体进行吸附，吸附后离心沉淀抗原-抗体复合物；最后将沉淀的复合物进行SDS-PAGE电泳，使得各蛋白质按相对分子质量大小分开，进行放射自显影，分析蛋白表达情况。

（三）免疫印迹（Western blot）

Western blot是20世纪70年代末、80年代初发展起来的一种结合了蛋白质凝胶电泳和固相免疫的蛋白质分析技术，兼备了凝胶电泳的高分辨力和固相免疫的特异敏感等特点，无需对蛋白质进行放射性标记，即可从混杂抗原中检测出微量特定抗原。其敏感度高，可检测出1～5ng的蛋白质，即被检蛋白占总蛋白的1/10即可检出。

免疫印迹技术包括三个步骤：

1. SDS-PAGE电泳　裂解细胞，进行蛋白质SDS-聚丙烯酰胺凝胶电泳，使蛋白质按相对分子质量大小分开；

2. 电转移　将凝胶上的蛋白质条带在电场力作用下转移至硝酸纤维素膜上；

3. 免疫显色　在硝酸纤维素膜上进行显色。先与一抗反应后再与相应酶标二抗反应并进行显色。

（四）ELISA（酶联免疫吸附法）

ELISA可用于抗原的定量分析。经上述特异性鉴定后，还应鉴定表达产物是否具有生物学活性及其活性高低，以判断是否具有实用价值。各种目的产物的性质不同，其生物学活性的鉴定方法也不同。如：肿瘤坏死因子α（TNFα）可通过测定其对小鼠L929细胞株的杀伤力来鉴定其活性；干扰素-α（IFNα）则通过测定其对人上皮细胞Wish细胞株的保护力来确定其活性。

八、原核生物表达系统的应用措施及其局限性

（一）外源基因在原核细胞中表达的应用措施

提高外源基因表达效率，可以从以下方面考虑：

1. 提高翻译水平　构建载体时，调整SD顺序与ATG之间的距离，或改变ATG与SD附近几个核苷酸，以消除可能出现的二级结构，提高翻译水平；也可以在DNA下游插入一转录终止子以稳定mRNA，提高翻译效能。

2. 选择合适宿主　不同菌株表达水平不同，不同载体适用于不同宿主。

3. 减轻宿主细胞代谢负荷，提高外源基因的表达水平 可以将宿主菌的生长和外源基因的表达分开（用 *PL* 启动子的温度诱导和 *lac* 启动子的化学诱导）；也可以将宿主菌的生长和表达质粒的复制分开（质粒 pCI101 转化宿主菌，在 25℃ 时宿主菌中质粒拷贝数仅 10 个，宿主大量生长，温度升至 37℃ 时，质粒大量复制）。这一方法是基因诱导表达的常用方法。

4. 提高表达蛋白稳定性，防止被宿主降解 产生融合蛋白或分泌蛋白均可达到此目的。另外，有些细菌突变株，失去或降低了蛋白酶的降解能力，可以用做受体：如 Lon-缺陷型菌株，缺乏合成蛋白酶所必须的肌苷（Lon）；再者，T4 噬菌体的 *Pin* 基因产物是细菌蛋白酶的抑制剂，将其克隆入载体可稳定基因产物。

（二）原核表达系统的局限性

1. 原核细胞中缺乏翻译后加工修饰系统，所产生的多肽不能进一步加工，比如糖基化、磷酸化，以及无法形成正确的蛋白质高级空间结构等，因而原核系统对很多真核蛋白都无法实现功能表达。

2. 原核表达外源蛋白常以包涵体形式存在，必须经过变性、复性处理，才能获得有生物学活性的蛋白，而复性的处理过程复杂，对相对分子质量大，尤其是分子内二硫键多的蛋白质，其表达效率非常低。

3. 在原核细胞中进行分泌表达，大多是将产物分泌至胞周间隙，很难实现真正的分泌出胞，其产量很低，而且容易出现信号肽不被切割，或不在特异位置上切割等问题。

4. 在原核细胞中表达的真核蛋白不稳定，易被细菌蛋白酶降解。

5. 原核细胞没有转录后加工系统，无法识别内含子，故原核细胞无法表达真核基因组 DNA。

6. 重组原核细胞在培养过程中产生的内毒素难以排除，污染表达产物，影响产品纯度。

（马春红）

第八章　真核生物表达系统

Chapter 8. Eukaryotic Expression System

真核生物基因表达系统(gene expression in eukaryotic cells)是指在真核细胞中表达外源基因的体系。真核细胞,不论是其基因结构,还是其表达调控都十分复杂,许多问题人们至今仍处在探讨之中。真核细胞基因工程(体外表达真核基因)起步较晚,1978年,Berg等采用SV40病毒做载体,在非洲绿猴肾组织培养细胞中成功表达了兔血红蛋白β多肽链,这是应用DNA体外重组技术,第一次成功地在真核细胞中表达另一真核基因,这在基因工程的理论和应用上都有重大意义。随着技术方法的突破,现在人类克隆成功的真核cDNA数目、种类都已极为繁复,而真核细胞表达系统的受体也已发展成为包括酵母、昆虫细胞、哺乳动物细胞等的一个大家族,真核细胞基因工程技术日趋成熟。关于真核基因在原核细胞中的表达,已在第七章中讨论,本章仅介绍基因在真核细胞中表达。

第一节　真核生物表达系统概述

一、真核生物表达系统的优势及必要性

真核生物表达系统由于真核细胞的基因结构及基因表达方面的特点而具有优越性:

1. 真核生物系统具有转录后加工系统,可识别并删除基因中的内含子,剪切加工为成熟mRNA;

2. 具备完善的翻译后加工系统,可进行糖基化、乙酰化、磷酸化等修饰,使蛋白质形成正确的天然构型,因而真核生物表达系统产生的蛋白质更接近天然状态,有利于对其功能、生物学活性的研究,也更具有价值;

3. 某些真核细胞可将基因表达产物直接分泌至细胞培养液中,简化了纯化工艺,降低了生产成本。

另外,由于真核细胞中基因表达的调控方式极其复杂,真核生物表达系统的研究有助于我们深入了解基因调控的机制。

二、真核生物基因结构及表达调控特点

(一) 真核生物基因结构特点

1. 真核生物中DNA极其丰富　真核细胞DNA是大肠杆菌DNA含量的数千倍(大肠杆菌基因组约4×10^6bp,哺乳类基因组在10^9bp数量级),这些DNA组成了真核生物的基因

组。要从如此丰富的基因组中克隆、筛选出某一目的基因就比较困难。克隆真核基因常用的方法是从细胞 mRNA 反转录合成相应的 cDNA。DNA 含量高、种类多,形成的 mRNA 也多,而相应的某一目的基因 mRNA 含量相对较少,考虑到真核生物各种组织细胞含有种类、数目完全相同的基因,但在各分化不同的细胞中表达不同的蛋白质,其 mRNA 转录有所不同,因而真核基因工程的第一个问题就是要选择易取材、易培养且含目的基因 mRNA 较丰富的细胞或组织作为目的基因的材料来源。

2. 真核基因的不连续性　真核生物基因的不连续性是真核生物基因结构上最重要的特点,即一个基因的编码序列(外显子,exon)被一个或几个非编码序列(内含子,intron)所分隔。由基因转录形成的 mRNA 前体必须经过剪切(splicing)、加工、切除内含子,才成为成熟的 mRNA,并由细胞核进入细胞质,这就增加了基因表达调控的环节。

3. 真核生物 mRNA 5′端具有帽子结构,3′ 端具有 polyA 尾巴　真核细胞 mRNA 5′ 端含有帽子结构:7m GPPN1mN2m…,3′ 端一般含有几十个至 200 个左右的多聚腺苷酸(polyA)尾巴(组蛋白除外),原核细胞中没有这种结构。这种差别导致了真核细胞 mRNA 半衰期较长,可达数小时甚至更长,因此可有足够时间提取 mRNA 并反转录合成 cDNA,使真核生物基因工程成为可能。

4. 真核生物基因组中含有大量重复序列　重复序列(repetitive sequences)即在基因组中多次重复出现的 DNA 序列。用复性动力学等实验表明有三类重复序列:①高度重复序列(high repetitive sequences)这类序列一般较短,长 10～300bp,在哺乳类基因组中重复 10^6 次左右,占基因组 DNA 序列总量的 10%～60%,人的基因组中这类序列约占 20%。②中度重复序列(moderate repetitive sequences),这类序列多数长 100～500bp,重复 10～10^5 次,占基因组的 10%～40%。例如哺乳类中含量最多的一种称为 Alu 的序列,长约 300bp,在哺乳类不同种属间相似,在基因组中重复 3～5×10^5 次,在人的基因组中约占 7%。在人的基因组中 18S/28S rRNA 基因重复 280 次,5S rRNA 基因重复 2000 次,tRNA 基因重复 1300 次,5 种组蛋白的基因串连成簇重复 30～40 次,这些基因都可归入中度重复序列范围。③单拷贝序列(single copy sequences),这类序列基本上不重复,占哺乳类基因组的 50%～80%,在人基因组中约占 65%。绝大多数真核生物为蛋白质编码的基因在单倍体基因组中都不重复,是单拷贝的基因。重复序列在真核细胞中行使各种功能:①编码一些功能十分重要、需要量极大的产物,如组蛋白、rRNA、tRNA 及某些糖代谢酶类等。这些重复序列满足了细胞的需要。②重复序列中含有特殊结构,参与基因表达调控:如 AT 含量丰富,易解链,有利于与调控基因复制的蛋白质因子结合,或者含有末端反向重复序列,可形成茎环结构。③与染色体构象、着丝点形成有关。

(二) 真核生物基因表达调控的特点

1. 真核生物基因表达的调控是多方面、多层次的。真核生物中 DNA 多与组蛋白结合形成复杂的高级结构,被包裹在核膜内,核外还有遗传成分(如线粒体 DNA 等),这就增加了基因表达调控的层次和复杂性。DNA 在核内转录后,经加工、切除,运输到核外才成为成熟的 mRNA,mRNA 在细胞质中与核糖体结合开始翻译。这一过程的每一步都影响基因的表达。

2. 一条成熟 mRNA 只能翻译出一条单一肽链，而不能形成原核生物的多顺反子 mRNA。

3. 真核生物中转录和翻译在时间上、空间上都是完全分开的，其功能相关基因分开，分别进行转录，而没有操纵子式结构。

（三）真核生物基因表达的调控模式

真核生物中基因表达的调控是十分复杂的，可以在转录前、转录、转录后、翻译和翻译后五个水平进行调控，其中转录水平的调控是最重要的一步。

1. 转录前（基因组水平）的调控　即基因结构上的改变，较稳定长久，包括基因丢失、基因重组、基因扩增，甲基化修饰和染色质结构改变。

2. 转录水平的调控　主要涉及三种因素：

（1）RNA 聚合酶：真核生物共有三种 RNA 聚合酶，分别合成不同类型的 RNA，其中 RNA 聚合酶Ⅱ转录合成 mRNA。

（2）顺式调控因子（*cis*-acting factor）：顺式调控因子是与结构基因串联的特定 DNA 序列，对基因转录的精确起始和转录活性有重要作用。主要包括起正性调控作用的顺式作用元件［启动子（promoter）、增强子（enhancer）］和起负性调控作用的元件——沉寂子（silencer）、加尾及转录终止信号。

1）启动子（promoter）：真核生物中的启动子（promoter）是位于基因 5′ 末端与转录起始有关的核苷酸序列。其作用特点为近距离（100bp）作用、具有方向性、空间位置较恒定。与原核生物相比，真核生物启动子更复杂、序列也更长。真核生物启动子间没有明显共同一致的序列，且真核生物启动基因转录需要多种蛋白质因子的相互协调作用，不同蛋白质因子又能与不同 DNA 序列相互作用，不同基因转录起始及其调控所需的蛋白因子不同，因而不同启动子序列也大不相同。真核生物启动子一般包括转录起始点及其上游约 100～200bp 序列，包含有若干具有独立功能的 DNA 序列元件，每个元件约长 7～30bp。最常见的哺乳类 RNA 聚合酶Ⅱ启动子中的元件序列见表 8-1。

表 8-1　哺乳类 RNA 聚合酶Ⅱ启动子中常见的元件

元件名称	共同序列	结合的蛋白因子		
		名称	相对分子质量	结合 DNA 长度
TATA box	TATAAAA	TBP	30 000	～10bp
GC box	GGGCGG	SP-1	105 000	～20bp
CAAT box	GGCCAATCT	CTF/NF1	60 000	～22bp
octamer	ATTTGCAT	Oct-1	76 000	～20bp
		Oct-2	53 000	～23bp
κB	GGGACTTTCC	NF-κB	44 000	～10bp
ATF	GTGACGT	AFT	?	20bp

真核生物启动子中的元件可以分为两种：①核心启动子元件（core promoter elements）指 RNA 聚合酶起始转录所必需的最小的 DNA 序列，包括转录起始点及其上游 －25/－30bp处的 TATA 盒。核心元件单独起作用时只能确定转录起始位点和产生基础水平的

转录。②上游启动子元件(upstream promoter elements)包括通常位于－70bp附近的CAAT盒和GC盒、以及距转录起始点更远的上游元件(图8-1)。这些元件与相应的蛋白因子结合能提高或改变转录效率。不同基因由不同的上游启动子元件组成,其位置也不相同,就使得不同的基因表达分别有不同的调控。

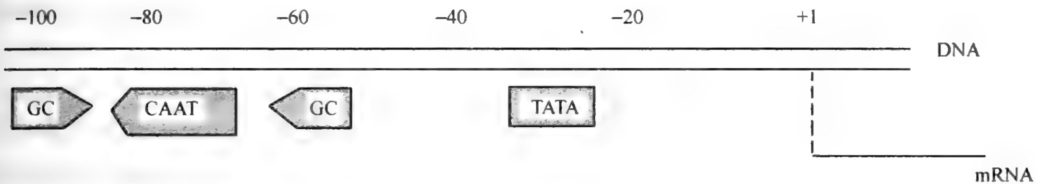


图8-1　真核基因启动子结构模式图

2) 增强子(enhancer):增强子(enhancer)是一类能促进转录活性的顺式调控元件,但它本身并不具备启动子活性。

增强子最早是在SV40病毒中发现的长约200bp的一段DNA,可使旁侧的基因转录提高100倍,其后在多种真核生物、甚至在原核生物中都发现了增强子。增强子通常由若干组件构成,长100～200bp,其核心组件常为8～12bp,可以单拷贝或多拷贝串连形式存在。

增强子作用有以下特点:无方向性、远距离作用、位置不恒定但有相位性(只有在DNA双螺旋的某一相位,活性较强)、无基因特异性、有组织特异性。

首先,增强子可以在任一位置提高同一条DNA链上基因的转录效率,可以远距离起作用,通常可距离1～4kb,而且在基因的上游或下游都能起作用。

其次,无方向性。增强子的作用与其序列的正反方向无关,将增强子方向倒置依然能起作用。

第三,增强子发挥作用依赖启动子的存在。没有启动子,增强子不能表现活性。但增强子对启动子没有严格的专一性,同一增强子可以影响不同类型启动子的转录。例如,当含有增强子的病毒基因组整合入宿主细胞基因组时,能增强整合区附近宿主某些基因的转录;当增强子随某些染色体段落移位时,也能提高移到的新位置周围基因的转录。

最后,增强子具有组织特异性。增强子的作用机制虽然还不明确,但与其他顺式调控元件一样,必须与特定的蛋白质因子结合后才能发挥增强转录的作用。许多增强子只在某些细胞或组织中表现活性,这种组织或细胞特异性是由这些细胞或组织中具有特异性蛋白质因子所决定的。

3) 衰减子(dehancer):衰减子(dehancer)或称沉寂子(silencer)是一类抑制基因转录的顺式调控因子,其作用方式与增强子相同:其作用可不受序列方向的影响,也能远距离发挥作用,并可对异源基因的表达起作用。

4) 转录终止信号:转录终止信号控制基因转录的结束。大多数真核基因其mRNA 3′端具polyA尾,其上游10～20bp处具有加尾信号(AATAA),ployA尾下游有G/T簇。加尾信号和G/T簇共同构成这些基因的转录终止信号。组蛋白和少数不具备polyA尾的基因的转录终止信号是一段能形成发夹结构的反向重复序列。

(3) 反式作用因子(*trans*-acting factor):是由位于不同或相同染色体上相距较远的基因

所编码的蛋白质因子,通过与顺式调控元件和 RNA 聚合酶的相互作用而调节基因转录活性。反式作用因子结构上包含一个与 DNA 结合的结构域和一个转录活化结构域,DNA 结合域靠它特有的结构与相应的 DNA 序列结合后,激活转录活化结构域,参与基因表达调控。

3. 转录后修饰　包括前体 mRNA 的剪切,内含子的去除,外显子拼接及 mRNA 的加尾和戴帽。加尾和戴帽均可增强 mRNA 稳定性,延长其寿命;另外,加尾帮助 mRNA 进入细胞质,戴帽则是核糖体结合和翻译起始所必需的。

4. 翻译水平的调控　主要涉及 mRNA、tRNA、核糖体和可溶性蛋白因子。其中反义 RNA/DNA 技术是目前使用较多的翻译水平调控措施。反义 RNA(antisense RNA)是一段含有与目的基因产生的 mRNA 碱基序列互补的小分子 RNA。反义 RNA 可通过与 mRNA 的结合,形成双链结构,影响 mRNA 的翻译。反义 RNA 对转录也有影响。天然反义 RNA 最初发现于原核生物,后来在真核生物中也有发现。目前人们在真核生物中模拟原核生物反义 RNA/DNA 作用机制,应用于抗病毒、抗寄生虫、抑制癌基因等研究,已取得很大进展。

5. 翻译后水平调控　蛋白质合成后,还必须经过一系列加工过程才能成为有生物活性的功能蛋白质,包括切除信号肽序列,进行磷酸化、乙酰化、糖基化等化学修饰等。

三、真核生物表达系统

真核生物表达系统中根据受体细胞的不同,其所应用的载体和表达策略各有不同,据此,可将真核生物表达系统分为三类:

1. 酵母表达系统　利用酵母表达载体,将外源基因在酵母细胞中表达;
2. 哺乳动物细胞表达系统　利用重组病毒载体,在哺乳动物细胞中表达外源基因;
3. 昆虫细胞表达系统　利用昆虫杆状病毒表达载体,在昆虫细胞中表达外源基因。

第二节　酵母表达系统

酵母表达系统中以酵母为受体细胞。酵母是单细胞真核生物,其遗传背景清楚,具有真核细胞的特点:可以识别内含子,对蛋白质进行多种翻译后修饰,有分泌功能,可对多种蛋白质进行胞外分泌表达,而其本身的分泌蛋白很少,因而纯化较方便。同时,酵母又具有原核细胞培养简单、生长周期短、容易操作的特点。此外,酵母具有较高的安全性,已被美国食品和医药管理局(FDA)确认为一种安全的生物。因而酵母是一种理想的基因工程受体细胞,酵母表达系统是一种很好的真核基因表达系统,有着广阔的发展前景。已有多种蛋白质在酵母细胞中成功表达,如人乙肝病毒表面抗原、核心抗原、干扰素、人表皮生长因子和胰岛素前体等。

一、酵母载体

(一) 酵母载体的组成与特点

酵母载体适用于在酵母细胞中复制、扩增和表达基因,分为克隆载体和表达载体,其主

要组成元件包括：

1. 遗传标记　所有酵母载体中都必须含有一定的遗传标记，以利于重组子的筛选、鉴定。酵母细胞虽可表达细菌抗生素，但却对大多数抗生素不敏感，因而多数细菌抗生素基因不能作为酵母细胞的遗传标记。酵母中常含有一些能与特定大肠杆菌变异株功能互补的一些基因，即这些基因相当于大肠杆菌的特定功能基因，这些基因常用作外源基因转化酵母时的遗传标记。

2. 用于酵母载体中的调控序列

(1) ARS(auto replication sequence)：酵母的自主复制序列，相当于复制起始区(点)，可启动基因在酵母中的复制。

(2) 2μm 质粒：酵母细胞中存在许多不同类型的内源性质粒，其中一种称为 2μm 质粒，它以高拷贝数存在于细胞核外，为小的环状 DNA，长约 2μm，可以独立复制并转录。

(3) 酵母启动子：在酵母表达载体中必须含有酵母启动子，以启动外源基因在酵母细胞中的表达。常用酵母启动子有：*ADH*(醇脱氢酶)基因、*PGK*(磷酸甘油激酶)基因、*PHO5*(酸性磷酸酶)基因、*GAP*(3-磷酸甘油醛脱氢酶)基因、*SUC* 蔗糖酶基因等。其中，磷酸甘油激酶(PGK)和磷酸甘油醛脱氢酶(GAP)在细胞内各占总蛋白的 5%左右，属于高含量蛋白质，因此，它们的启动子为强启动子。

(4) 酵母着丝粒区段(cEN)：增加转化子稳定性。

3. 酵母载体均是由细菌质粒 DNA 与酵母 DNA 重组的穿梭质粒(shuttle vector)，既可以在原核生物中复制增殖，又可在真核酵母中复制和表达。

(二) 酵母载体的分类

酵母载体可分为五类，其结构和构建过程如图 8-2。

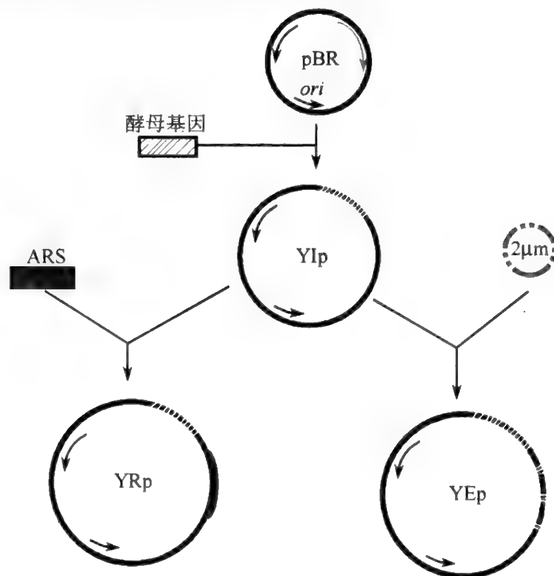


图 8-2　酵母载体的种类和构建过程

1. 酵母整合型质粒(YIp,yeast integrating plasmid) 在细菌质粒中引入酵母的遗传标记基因,即成为 YIp。YIp 中不含酵母复制起始位点,而是通过同源重组,整合入酵母染色体,形成稳定的转化子,并随染色体复制而复制。YIp 转化率极低(1～10 转化子/μg DNA)。

2. 酵母复制型质粒(YRp,yeast replicating plasmid) 由细菌质粒和带有复制子(*ARS*1 基因)及遗传标记的酵母基因组成,可在酵母细胞中自主复制,其转化频率比 YIp 高 100 倍以上,拷贝数也多,但 YRp 不稳定,即使在有选择压力下也容易丢失。

3. 酵母附加子型载体(YEp,yeast episomal plasmid) 由细菌质粒 DNA、酵母标记基因和酵母 2μm DNA 组成。酵母 2μm DNA 是酵母中类似于质粒的一种游离于核外的 DNA,可独立复制。因而由 2μm DNA 组成的 YEp 型载体可以在核外存在并独立复制,其转化频率、拷贝数及稳定性与 YRp 相似。

上述三种酵母载体均不含酵母启动子,其中的原核启动子没有能力启动真核基因在酵母中的表达,因而不适用于在酵母中表达外源基因。

4. 酵母表达质粒 在上述酵母穿梭质粒基础上添加酵母启动子,即可在酵母中表达基因。若在酵母表达型质粒中加入信号肽序列,则可以实现外源基因的分泌表达。

5. 酵母人工染色体(yeast artificial chromosome,YAC) 由酵母染色体和 2μm DNA 的复制起始区、着丝粒序列和四膜虫端粒 DNA 组成。YAC 在细胞有丝分裂和减数分裂中具有天然染色体的稳定性。YAC 允许插入 300kb 的外源片段,甚至长达 1000kb 以上的片段也能插入并稳定转化,有利于对大片段克隆 DNA 的表达及蛋白功能进行研究,也使建立基因组文库的工作量减少。图 8-3 显示了以 YAC 克隆基因的基本过程。

2μm 质粒,后来研究人员相继又构建了一系列酵母的克隆载体和表达载体。另外,啤酒酵母具有较高的安全性,已被 FDA 确认为一种安全的生物,以其为受体表达的产品不需要进行大量的宿主安全性实验。

利用啤酒酵母已成功表达了多种蛋白质,但是啤酒酵母作为受体细胞,有不可避免的缺点:表达量较低,大量表达时,易发生质粒的丢失;重组蛋白经常出现过量糖基化;分泌表达时,有时会将产物留在壁膜间隙中。

毕赤酵母系统是最近发展起来的新型表达系统,与啤酒酵母比较,具有下列优点:

(1) 表达量高:利用强启动子(如:醇氧化酶 *AOX* 基因的启动子),表达量较其他表达系统高 10 倍至 100 倍以上。以毕赤酵母表达的破伤风毒素 C,其表达量可达 12g/L。

(2) 稳定性好:毕赤酵母的表达载体为整合型载体,可通过同源重组整合至酵母染色体中,形成稳定的转化子。

(3) 高分泌:以啤酒酵母的 α 因子信号肽引入毕赤酵母表达载体,可获得很好的分泌效果,人血白蛋白分泌表达可达 10g/L。

二、外源基因在酵母中的表达策略

1. 酵母的偏爱密码子 不同生物中密码子的使用频率不同,同一种生物细胞对同一氨基酸的不同密码子使用频率也不相同。在酵母中表达外源基因,自然要选用酵母偏爱的密

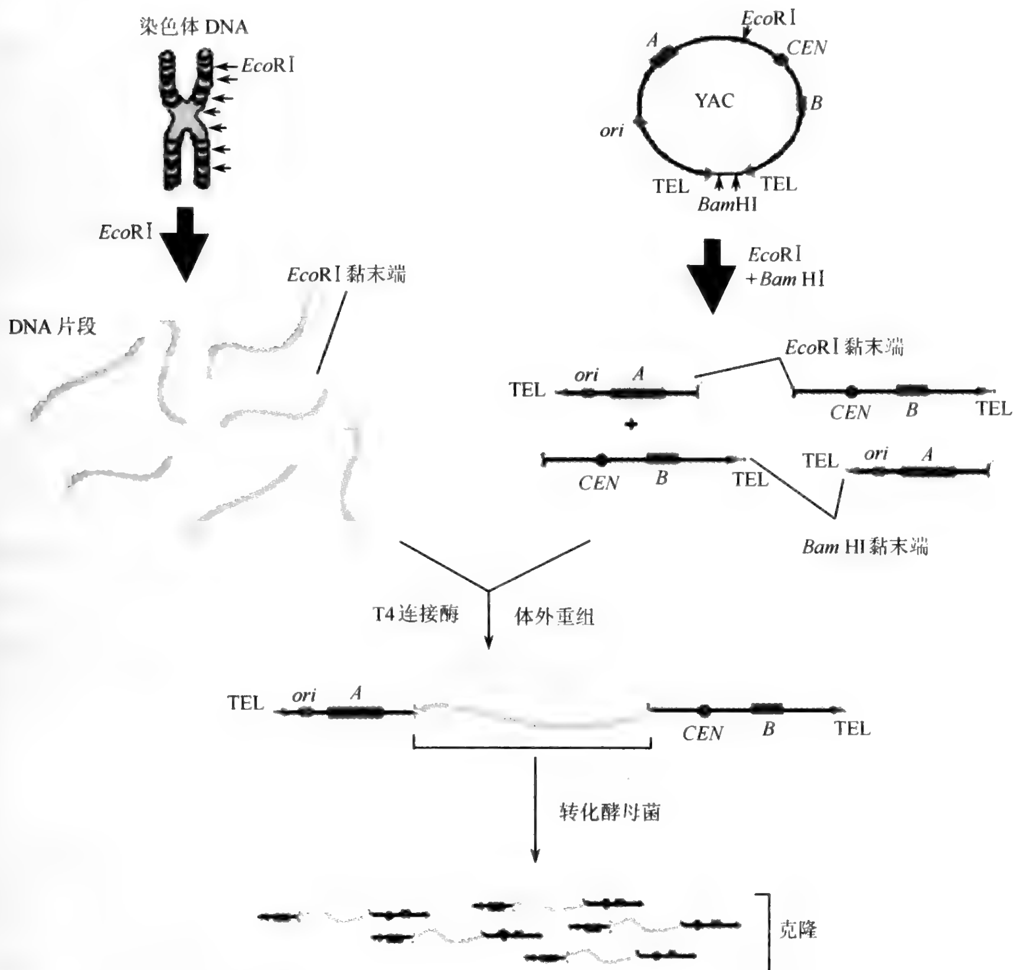


图 8-3 YAC 克隆基因的流程

码子,但同时要考虑基因转录产物 mRNA 的二级结构对表达蛋白的影响。

2. 外源基因的表达形式 外源基因在酵母中有两种表达形式:

(1) 直接表达外源基因:即将外源蛋白表达在细胞质中,而不分泌出胞。酵母菌表达外源基因,多为胞内表达,其中,乙肝病毒表面抗原在酵母菌中的表达就是一个成功的例子(图 8-4)。

(2) 分泌表达:分泌表达可以避免产物被降解,同时有利于产物纯化,并帮助产物进行糖基化修饰,是一种很好的表达形式。在酵母细胞中利用外源性的信号肽很难分泌表达。然而利用酵母细胞中天然存在的分泌型蛋白质的信号肽,可以构建酵母分泌型表达载体,实现外源基因的分泌表达。最常用的酵母分泌信号肽序列是啤酒酵母 α 因子的信号肽。pGAPZα(图 8-5)是利用 α 因子进行分泌表达的一种新型毕赤酵母表达载体。影响酵母细胞分泌表达的因素除信号肽外,还有许多其他因素,比如:基因剂量、营养型及表达产物的产量等。

3. 表达策略 为实现外源基因在酵母中的高效表达,可采取如下措施:

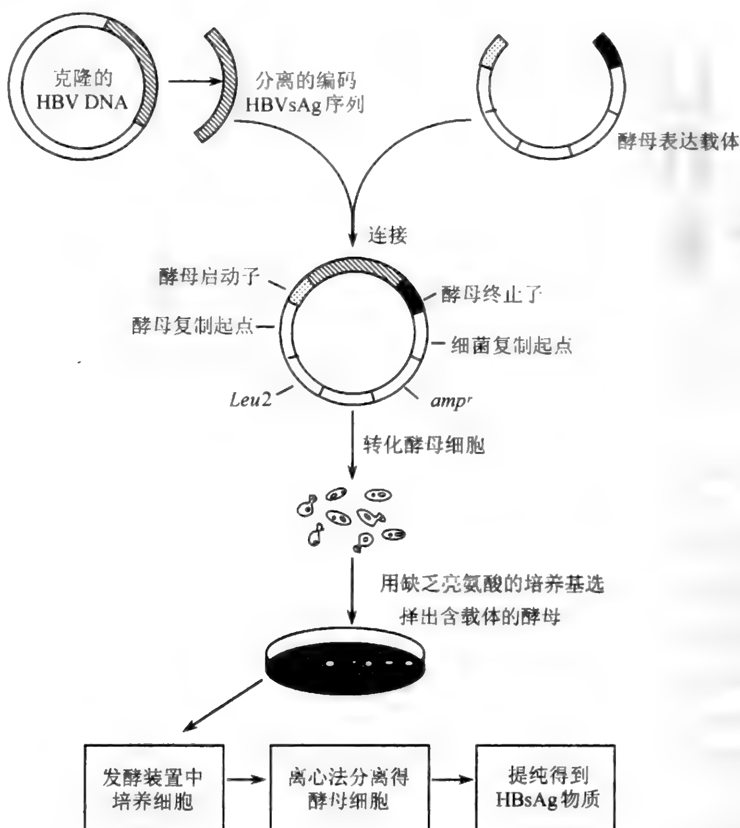


图 8-4 酵母细胞表达乙肝病毒表面抗原流程图

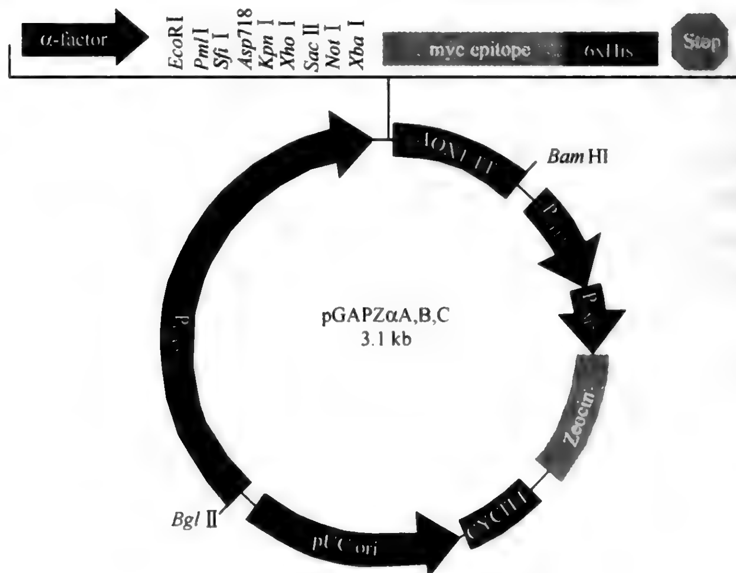


图 8-5 pGAPZα 物理图谱

(1) 利用诱导型强启动子:毕赤酵母表达载体中常用的醇氧化酶 *AOX* 启动子是一种甲醇诱导型启动子,可利用甲醇诱导外源基因的表达。

(2) 提高整合型表达载体在细胞中的整合拷贝数:可利用载体中提供的抗生素遗传标记,以抗生素加压筛选含有多拷贝的转化子。

(3) 控制蛋白酶降解活性:采用蛋白酶缺陷型宿主菌,降低发酵培养基的 pH,或在培养基中补加氨基酸或多肽,均可避免产物被降解。此外,分泌表达也是一种避免产物被降解的良好策略。

第三节　哺乳动物细胞表达系统

哺乳动物细胞表达系统是指在哺乳动物细胞中表达目的基因,或采用某种方式,将目的基因导入哺乳动物细胞,改变受体细胞的某种状态。

哺乳动物细胞是自然界中最高级的真核细胞,以其为受体细胞进行基因表达,可对表达产物进行完全的翻译后修饰,包括二硫键形成、糖基化、磷酸化、寡聚体形成及蛋白酶对蛋白质的特定切割。哺乳动物细胞还可表达有功能的膜蛋白受体细胞,这为研究多种膜受体的生物学功能提供了必要工具。此外,哺乳动物细胞表达系统是进行基因治疗的惟一受体系统。因而,哺乳动物细胞表达系统有着其他表达系统不可替代的作用。

一、哺乳动物细胞表达系统中的常用遗传标记

1. 胸苷激酶(*tk*)基因——HAT 选择法　正常细胞中二氢叶酸还原酶将二氢叶酸还原为四氢叶酸,参与合成 dTTP、dATP、dCTP。当培养基中加入二氢叶酸类似物氨基蝶呤时,二氢叶酸还原酶失活,上述途径受阻。由于正常细胞中含有 *tk* 基因,*tk* 基因可将胸苷转化为胸苷一磷酸,进而继续合成 dTTP,满足细胞需求;若同时在培养基补加次黄嘌呤作为 dATP、dCTP 合成的替补途径,则细胞可以生存。因而在 HAT 培养基(正常培养基中添加氨基蝶呤、次黄嘌呤及胸苷)中 tk^+ 细胞可以存活,而 tk^- 细胞因为缺乏 *tk* 基因,不能合成 dTTP,不能在 HAT 培养中生存。

因此,以 tk^- 细胞株为受体细胞,以含次黄嘌呤(hypoxanthine)、氨基蝶呤(aminopterin)和胸苷(thymidine)的 HAT 选择培养基进行选择培养,可筛选重组有 *tk* 基因的 tk^- 细胞。

2. 二氢叶酸还原酶(dihydrofolate reductase, dhfr)基因　dhfr 是真核细胞合成胸腺嘧啶过程中重要的酶,可催化二氢叶酸还原为四氢叶酸,后者为合成胸腺嘧啶提供甲基。这一过程可被叶酸类似物氨基蝶呤和甲氨蝶呤(methotrexate,MTX)竞争性抑制。$dhfr^-$ 细胞在普通培养基中无法成活,但转入 *dhfr* 基因后可生存。因而,以 $dhfr^-$ 细胞为受体,可选择转入 *dhfr* 基因的细胞。

dhfr 为选择标记对受体细胞有严格的限制,但该选择系统的优点是:若培养基中加入甲氨蝶呤(MTX),并逐渐增加浓度,使细胞逐渐产生抗性,可以使转入的 *dhfr* 基因和外源基因明显扩增,外源基因表达量增加。

3. 新霉素抗性基因(*neo*)　此标记基因适合在含有新霉素类似物 G418(gentamycin)的

选择培养基中选择 *neo* 的细胞。其原理为:虽然新霉素只对原核细胞有毒性而对真核细胞无毒性,但其类似物 G418 却对真核、原核细胞均有毒性,故一般真核细胞在含 G418 培养基中不生长。*neo* 基因编码的磷酸转移酶可以使 G418 失活,因而真核细胞中转入 *neo* 基因后可以在含 G418 的培养基中生长。

以新霉素抗性基因为选择标记适用于所有真核细胞,对受体细胞无特殊要求,是一种应用极为广泛的真核标记基因。

转化后的重组子经上述筛选标志初选后,还应进行特异性筛选:可进行原位杂交或提取重组子核酸进行酶切电泳分析,有必要时可进行核酸序列分析和表达产物检测。

二、哺乳动物细胞表达载体

在哺乳动物细胞中表达外源基因,真核载体是必不可少的工具,也是成功的关键。所谓真核细胞载体即可携带外源基因进入真核细胞,并在其中复制、增殖、表达的工具。根据功能,可以分为真核细胞克隆载体和真核细胞表达载体。真核细胞克隆载体指适用于外源基因在真核细胞中扩增的载体工具。因为在真核细胞中扩增操作复杂、周期长、成本高,所以一般都在真核细胞克隆载体的基础上进一步加工,加入真核启动子,增强子等元件,使外源基因可以在真核细胞中表达,这就是真核表达载体(适用于外源基因在真核细胞中的表达)。另外,受体细胞不同,适用的载体也不同,这些载体又有一些特殊名称,其中,适用于哺乳动物细胞的载体统称为哺乳动物细胞表达载体。

(一) 哺乳动物细胞表达载体的主要组成元件

现在经常应用的哺乳动物表达载体均为穿梭载体,既包含原核序列,以利于在原核细胞中扩增,又包含能在真核细胞中表达的转录单位(由非编码序列及选择标志组成,常将已清楚的几种不同病毒及哺乳动物基因融合在一起)。哺乳动物表达载体主要包括以下元件:

1. 启动子　启动子的转录效率因细胞而异,因此要根据宿主细胞的类型选择不同的启动子。真核表达载体中常用的启动子包括病毒来源的启动子(如:绿猴空泡病毒 40,simian vacuolating virus40,SV40)启动子、巨细胞病毒(cytomegalovirus, CMV)启动子、肉瘤病毒(Rous sarcoma virus, RSV)启动子、反转录病毒长末端重复序列(long terminal repeat, LTR)等)和细胞来源的启动子[如热休克蛋白(heat shock protein,HSP)启动子、肌酸激酶 MCK 启动子等]。大多数真核启动子、增强子具有较强特异性,故应根据宿主加以选择;但有些来自于病毒的增强子宿主范围较广泛,如 SV40 早期增强子,可对多种哺乳动物型别细胞启动子有作用。

2. 增强子　许多来源于病毒的增强子具有广泛的宿主范围,在不同种属细胞中都有极强的促进转录能力,这些序列已被广泛用于真核表达载体中,常用的有 SV40、RSV、CMV、LTR 增强子。

3. 筛选标记　克隆在真核载体中的重组子,由于真核基因一般没有特殊标志,故必须在真核载体中填加筛选标志。常用的标志基因详见前述。

4. 转录终止信号和多聚腺苷酸化信号　为使转录后的 mRNA 能有效进行切割和多聚

腺苷酸化，在真核表达载体中必须含有转录终止信号和多聚腺苷酸化信号，最常用的 polyA 信号是来自 SV40 的一段 237bp 的 *Bam*H I-*Bcl* I 限制性酶切片段。另外，真核基因 3′非编码区的一些序列可影响 mRNA 稳定性，故有必要切除。

5. 剪接信号　并非所有 cDNA 都需要剪接信号，但一般来说，内含子不会降低而是提高表达效率，故可采取在哺乳动物基因转录单位中带有剪接供体和受体位点的载体。

6. 复制元件　载体中的复制元件可提高外源基因表达水平。如 SV40 复制子，可使载体在相应宿主中大量复制，达极高拷贝数（3～4 天后达 10^4/细胞），用于瞬时而大量表达蛋白。

7. 原核序列　为了能方便地获得大量的重组 DNA，哺乳动物表达载体通常含有一段原核序列，其中包括：能在大肠杆菌中复制的复制子、便于在原核细胞中筛选重组子的抗性基因和多克隆位点。

（二）哺乳动物细胞表达载体的种类

哺乳动物表达载体分为病毒型载体和非病毒型载体两大类。

1. 病毒型载体　病毒是在漫长的自然进化过程中存活下来的没有细胞结构的最小、最简单的生命寄生形式。它们通常可以高效率地进入特定类型的细胞，表达自身蛋白并产生新的病毒粒子。因此，病毒是首先被改造作为真核表达的载体。原则上讲，所有的病毒通过一系列的处理（如删除与致癌、致毒和复制相关的基因片段，在合适的位置插入外源基因），均可改造成为基因传递的工具。

目前常用的病毒载体分为两类：整合型和游离型。其中整合型载体整合入染色体，并随染色体复制而复制，可持续表达外源基因，但存在插入诱变的危险，主要包括 SV40 载体、反转录病毒载体。游离型载体不整合入染色体，而是游离于染色体外，以病毒颗粒的形式独立复制，或在辅助病毒的帮助下复制，没有插入诱变的危险，安全性高，但一般为瞬时表达。主要包括：腺病毒载体、痘苗病毒载体等。

（1）SV40 病毒载体：SV40 DNA 是一个小的、共价闭合双链 DNA，长约 5.2kb，其序列及物理图谱和各功能区段已清楚。SV40 基因组分早期区和晚期区，早期区基因是病毒溶源生长所必需，可产生 t 抗原和 T 抗原，晚期区则产生病毒衣壳蛋白，是包装病毒颗粒所必需的。

SV40 病毒载体是发展最早的经典的病毒载体，其构建方法有两种：一种是将外源基因取代早期区或晚期区，然后与辅助病毒（只包含早期区或晚期区）共同感染宿主，即可获得含外源基因的子代病毒颗粒，另一种是仅保存病毒基因组中与复制相关的序列，并与质粒 DNA 重组，构建穿梭载体。

$pSVK_3$ 质粒（图 8-6）是广泛适用于多种哺乳动物细胞的多功能表达载体。由以下元件构成：①质粒复制起始点 *ColEI*：质粒可在细菌中复制扩增；②丝状噬菌体复制起始点 *f*1：在辅助噬菌体 $M_{13}KO_7$ 帮助下可产生单链噬菌体；③SV40 的复制起始点、早期启动子、mRNA拼接信号和接尾信号：质粒可在真核细胞中复制并表达基因；④多克隆位点；⑤T7RNA聚合酶启动子：可在体外转录插入片段的 DNA。

（2）反转录病毒（RV）载体：反转录病毒属于正链 RNA 病毒，可高效地感染许多类型的

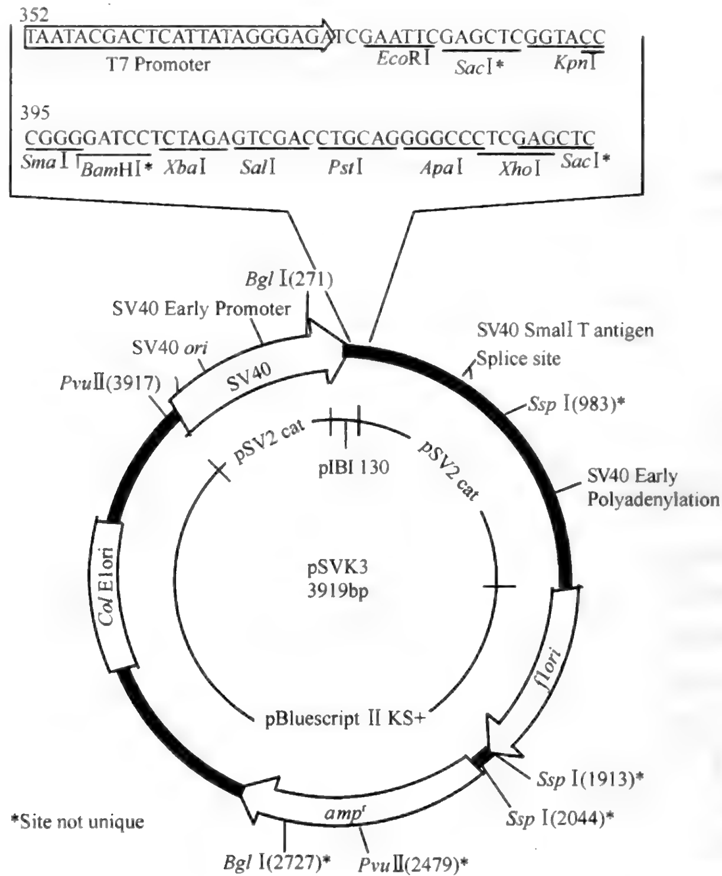


图 8-6 $pSVK_3$ 质粒

宿主细胞,并稳定整合到宿主细胞基因组中。它是最先被改造且应用最为广泛的基因治疗载体。

反转录病毒基因组大约 10kb,含有三个最重要的基因:*gag*(编码核心蛋白)、*pol*(编码反转录酶)和 *env*(编码病毒外膜蛋白)。两端存在长末端重复区(LTRS)用于介导病毒的整合。*env* 基因中含有病毒包装所必需的序列(包装信号)。

常用的反转录病毒载体多由 Moloney 鼠白血病病毒(MoMuLV)改造而来。此类病毒为嗜双性病毒,既可感染鼠细胞亦可感染人细胞。反转录病毒载体的基本成分包括:必需的病毒基因组分(如 LTRS、病毒的包装识别信号、翻译所需的剪接识别位点及 poly A 尾等)、标记基因和其他质粒必要成分。

反转录病毒载体均为复制缺陷型反转录病毒载体,其中缺失反转录病毒的 *gag*、*pol* 和 *env* 等基因,因而不能形成病毒颗粒,必须借助包装细胞提供病毒蛋白,才能包装产生子代重组病毒。这种子代重组病毒也仅仅具备一次感染性,从而避免了在人体细胞间的扩散感染,也大大降低了病毒本身的致癌性与致病性。所谓包装细胞即将缺失包装信号的反转录病毒 DNA 导入细胞即构建成包装细胞系,可为反转录病毒载体提供结构蛋白,使之包装形

成子代病毒颗粒。其中包装细胞含有反转录病毒的包装蛋白(满足穿梭载体需求),但缺失包装信号等调控元件,同样不能独自产生完整病毒颗粒。病毒载体转染进入包装细胞后,载体与包装细胞作用互补共同完成病毒包装,形成的病毒颗粒感染目的细胞后,由于目的细胞不含有编码病毒蛋白的基因,病毒 DNA 不能包装形成完整病毒颗粒,避免了扩散。

基于上述设想,人们建立了第一代包装细胞系 Φ2,Φ2 是单向性包装细胞,包装形成的病毒颗粒仅感染啮齿类动物。由于 Φ2 缺失包装信号,其反转录与整合信号完整,因而当载体转入后,易产生辅助病毒。辅助病毒是为缺陷型病毒提供包装所必需基因的病毒。若辅助病毒的基因组一起被包装,则产生的病毒颗粒进入受体细胞后仍可进一步感染其他细胞。第二代包装细胞的代表 PA317 是双向性包装(子代颗粒除感染啮齿类动物外,还可感染灵长类动物)。PA317 中除缺乏包装信号,还有 5′ LTR 端的缺失,且其 3′ LTR 被 SV40 的多聚腺苷酸位点取代,因而其产生辅助病毒的可能性大大减少。第三代包装细胞(ΦCRE、ΦCRIP)进一步在结构基因内引入突变与缺失,使得产生辅助病毒的可能性进一步降低。

反转录表达载体作为基因载体具有明显优势:宿主范围广泛,存在于几乎所有的人类细胞上,因而对细胞的感染谱广,特别是对分裂细胞感染效率较高;可主动感染分裂细胞;可整合到宿主染色体中,在细胞内长期存在而不杀死宿主细胞,表达时间较长,是当今基因治疗中常用载体。但反转录病毒载体也存在不可克服的缺点:仅整合到处于增殖状态的细胞,对静止细胞无感染能力;允许插入的外源基因有严格的限制,一般小于 8kb;携带外源基因随机整合人人体组织细胞,有诱变危险,且组织或细胞选择性不高。

(3) 腺病毒(AdV)载体:腺病毒是无包膜的线性双链 DNA 病毒,在自然界分布广泛,其基因组长约 36kb,两端各有一个反向末端重复区(ITR),ITR 内侧为病毒包装信号。基因组上分布着四个早期转录单位(*E*1、*E*2、*E*3、*E*4)承担调节功能,和一个晚期转录单位,负责结构蛋白的编码。

第一代腺病毒载体中腺病毒基因组上的 *E1A* 和 *E1B* 基因已被目的基因所取代,因而在普通体内缺乏复制能力,仅在染色体上已整合有 *E1A* 和 *E1B* 的包装细胞(如 293 细胞)中复制,缺陷型的重组腺病毒才能包装成完整的腺病毒颗粒,具有较好的安全性。但此类型载体可引发机体产生强烈的炎症反应和免疫反应,且表达外源基因时间短。第二代腺病毒载体在第一代载体的基础上,缺失 *E2A* 或 *E*4 基因,产生较弱的免疫反应,其载体容量和安全性方面较第一代载体改进许多。第三代腺病毒载体缺失了全部的(无病毒载体,gutless vector)或大部分腺病毒基因(微型腺病毒载体,mini-Ad),仅保留 ITR 和包装信号序列。第三代腺病毒载体最大可插入 35kb 的基因,病毒蛋白表达引起的细胞免疫反应进一步减少;载体中引入核基质附着区基因可使得外源基因保持长期表达,并增加了载体的稳定性。

腺病毒载体其生物学特性有:不整合到染色体中;外源基因表达水平高;表达时间较短;免疫原性强。

与反转录病毒载体相比,腺病毒载体具有显著优点:①基因组大,因而可插入大片段外源基因(最多可达 35kb);②可高效地转染不同类型的人体组织细胞(体外实验通常接近 100%转染效率);③可感染分裂和非分裂细胞;④病毒滴度高,可达 10^{11} pfu/ml;⑤可原位感染组织,如肺等;⑥进入细胞内并不整合到宿主细胞基因组,仅瞬间表达,因而安全性较高。腺病毒载体的缺点主要表现为:免疫原性强,可刺激机体产生免疫反应;重组、纯化费时;几

乎可以感染所有细胞，而缺乏特异性。

(4) 腺相关病毒(AAV)载体：腺相关病毒属于非致病的微小病毒科家族的成员。只有依赖于辅助病毒(如腺病毒)才能增殖。此类病毒基因组很小，如Ⅱ型腺相关病毒是由4681个核苷酸组成的单链DNA；包含两个基因：*rep* 基因(编码负责调节病毒复制、结构基因表达和整合到宿主基因组中的蛋白)；*cap* 基因(编码衣壳结构蛋白)，基因组的一个末端存在一个145bp的末端重复区(terminal repeat)。

AAV是目前基因治疗载体系统研究的热点之一。其生物学特点有：基因组小而易于操作；可感染分裂和非分裂细胞；当辅助病毒不存在时，AAV能整合到宿主细胞基因组的特定区域(如人的第19号染色体短臂上)，但构建的载体实际应用时则很难达到定向整合；无致病性、安全性好；免疫原性弱，实验证明其可以有效地转染脑、骨骼肌、肝脏等许多类型的细胞。缺点是装载外源基因容量有限(不超过4.9kb)，缺少高效的包装细胞，制备过程复杂，制备滴度低($<10^4$ 病毒颗粒/ml)，可能出现免疫排斥反应。

(5) 痘苗病毒载体：痘苗病毒是动物病毒中最大的一种，其基因组也非常大，且与天花病毒有密切亲缘关系，长期以来被用做预防天花的疫苗。

痘苗病毒的庞大基因组中含有一个大的复制非必需区，可供外源基因插入。痘苗病毒载体一般包括：痘苗病毒调控顺序、外源基因插入位点和选择标记。因为痘苗病毒复制非必需区很大，因而可能允许同时插入几种基因，构建多价疫苗。

2. 非病毒载体　虽然病毒载体作为基因传递的工具广泛被采纳，但它们仍存在着不少的局限性：病毒性载体均可以诱导机体产生某种程度的免疫反应；存在着插入突变的风险；载体容量有限；制备过程复杂等。

非病毒载体中含有原核组成元件和真核启动子元件，因而是一类穿梭载体，既可以在原核细胞中复制增殖，又可以在真核细胞中转录表达。非病毒载体不具备形成病毒颗粒的能力，不能通过包装，形成有感染性的活病毒主动感染受体细胞，而是需要借助一定的手段进入细胞。

与病毒载体相比，非病毒载体则具有无传染性、不限制载体容量，可大量制备等优点。

pCDNA3 (图8-7) 是目前常用的非病毒型表达载体，其主要组成元件包括：*CMV* 启动子、多克隆位点、*amp* 抗性基因和新霉素抗性基因、大肠杆菌复制子、*f*1 丝状噬菌体复制起始点、*SV*40 启动子、增强子和polyA加尾信号。

三、哺乳动物细胞基因转移方式

成功构建含有目的基因的重组表达载体后，必须要将重组子导入到受体细胞。哺乳动物细胞基因转移方法较多，可分为物理、化学和生物三大类。前两种是非病毒方法，包括磷酸钙沉淀法、DEAE-Dextran法、微量注射法，此外还有电穿孔法、裸露DNA直接注射、脂质体共转化和受体介导的基因转移等。第三种方法为病毒方法，指病毒载体重组后，通过体外包装，形成携带目的基因的病毒颗粒，主动感染靶细胞并表达出目的基因的方法。

1. 磷酸钙沉淀法　首先由Granhan和Vander EB建立，主要用于贴壁和悬浮细胞的转染。磷酸钙沉淀法的关键是以$CaCl_2$和磷酸盐形成合适颗粒的磷酸钙沉淀，沉淀形成后将

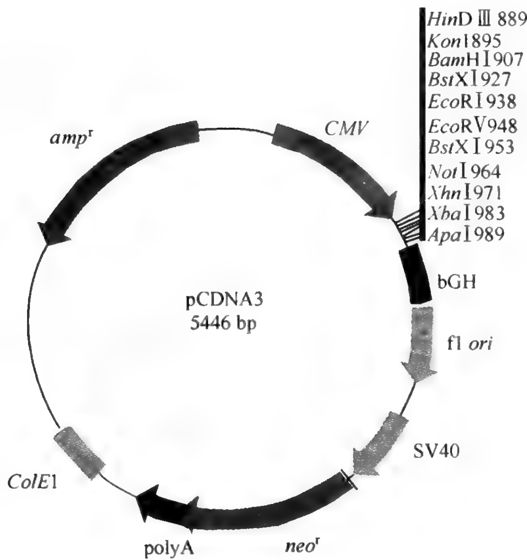


图 8-7　pCDNA3 物理图谱

DNA 与沉淀混合并加入对数生长期细胞，由于磷酸钙沉淀使细胞通透性增高，DNA 随之进入细胞。

2. DEAE-dextran 法　最早由 Pagano 建立于 1969 年，现已广泛用于转染技术，尤为暂时表达。该法主要是应用 DEAE-dextran 对细胞的毒性作用，将 DNA 和 DEAE-dextran 无血清培养基一起加到受体细胞上，在 DEAE-dextran 的作用下，DNA 导入受体细胞。特点是：DNA 用量少、简便快速，影响因素少；同时使用二磷酸氯喹处理细胞，可防止转入的 DNA 被降解，大大提高了转染效率。

3. 电穿孔法　根据细胞暴露于高电压时，细胞膜可被电脉冲瞬间穿孔，从而使存在于细胞悬液中的外源 DNA 进入细胞设计的高效基因转移方法。电穿孔法转染效率很高，但细胞受损严重。

4. 裸 DNA 注射　即把携带外源基因的质粒直接注射或通过基因枪导入到组织细胞中，此方法主要应用于 DNA 疫苗（外源基因为病原体的有功能的基因片段），可激发机体的细胞免疫和体液免疫。

5. 脂质体法　脂质体（liposome）是由脂质双分子层组成的颗粒，可介导基因穿过细胞膜。最常用的脂质体为阳离子脂质体（或称阳性脂质体），主要由带正电荷的脂类和中性辅助脂类等摩尔混合。带阳性电荷的脂质体与带阴性电荷的 DNA 之间可以有效地形成复合物，此复合物通过内吞作用可进入细胞中。其作用机制如图 8-8 所示。

脂质体介导的基因转染有以下特点：①脂质体与基因的复合过程比较容易；②脂质体与细胞膜融合将目的基因导入细胞后，脂质体即被降解，无毒，无免疫原性；③DNA 或 RNA 可得到保护，不至于灭活或被核酸酶降解；④转染过程方便易行，重现性好。

脂质体介导的基因转移方法具有简便、安全、可携带较大 DNA 分子等优点，已在多种基因治疗实验中成功应用，是目前基因治疗临床研究方案中，除反转录病毒载体以外应用最

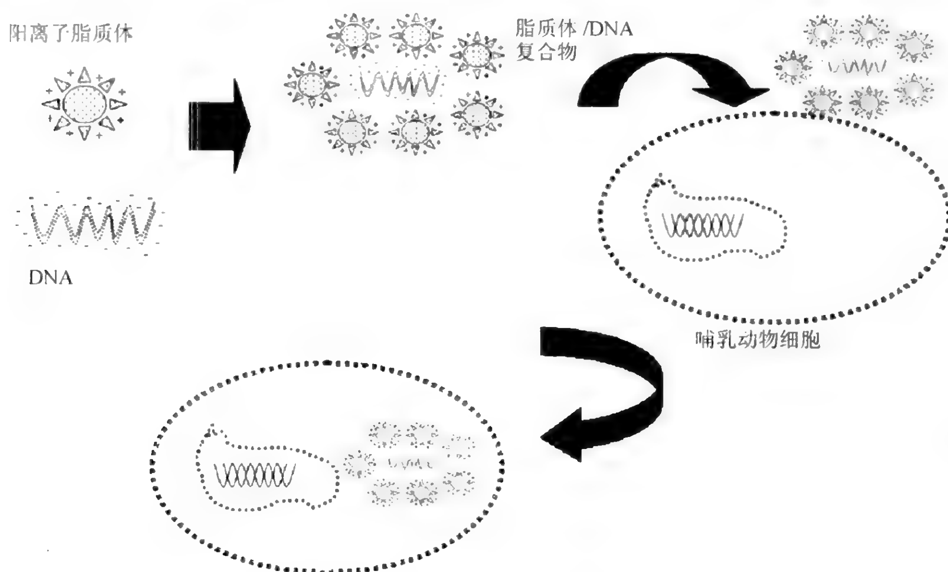


图 8-8　阳离子脂质体基因转移作用机制

多的基因传递方法(>20%)。但脂质体介导的转染效率低,尤其是无法解决细胞特异性导入问题。

6. 受体介导的基因转移　受体介导的基因转移是近年发展的新技术,它利用细胞表面存在的特异性受体介导的内吞作用,将偶联了受体特异性配体的 DNA 结合分子(如多聚赖氨酸)与 DNA 结合,形成可溶性复合物,通过内在化实现基因定向转移。

受体介导的基因转移系统中,外源基因被内吞后,形成吞噬溶酶体,常导致部分外源基因及其表达产物被降解而影响表达率。如何使得外源基因稳定地从吞噬溶酶体中释放,是提高受体介导基因转移系统转移效率的关键。1998 年,顾建人、田培坤等利用流感病毒部分功能域可使内吞小体释放其内吞物,从而逃避溶酶体降解的特点,建立含内吞小体释放寡肽的新型细胞受体为靶向的基因导入系统,显示了其在基因治疗中的优势和潜在应用价值。

除细胞表面存在的特异性受体介导外,也可以将 DNA 与单克隆抗体或病毒胞膜蛋白共价结合,而介导外源基因导入到某一类型的细胞中。

第四节　昆虫杆状病毒表达系统

杆状病毒(Baculovirus)是许多农林业昆虫的天敌,可作为微生物杀虫剂服务于人类。以杆状病毒作为外源基因表达载体,具有极其广泛的应用。1987 年统计,国际上已有超过 150 所大学、研究机构或公司的实验室从事这一研究,而且已表达几十种真核基因。其中,第一个为美国食品和药物管理局批准上人体试验,拟用于预防艾滋病的基因工程疫苗,就是应用杆状病毒表达载体系统表达生产的。

一、昆虫杆状病毒载体

(一) 杆状病毒基因组特点

杆状病毒基因组庞大，约 100～150kb，其中包含一些与病毒复制无关的非必需区。最早发现，也是应用最多的一个病毒复制非必需区是病毒的多角体编码基因(*Ocu*)。多角体蛋白是病毒复制晚期用来包埋病毒粒子的蛋白质成分，在光镜下明显可见，其基因启动子强大，可以启动外源基因高效表达。另外，后来发现病毒复制最晚期表达一种 10kD 的蛋白质，也是复制非必需区，其 P10 启动子也逐渐应用于杆状病毒载体。

(二) 转移载体

由于杆状病毒基因组庞大，难以在体外拼接，必须借助转移载体进行体外基因操作。

1. 转移载体(transfer vector)　指同时含有原核序列(复制起始信号、抗性基因)和病毒基因组序列的载体，除可携带外源基因在原核中复制、扩增外，还可通过细胞内同源重组，将外源基因插入病毒基因组，构建重组病毒。

2. 转移载体的基本元件(图 8-9)

(1) 原核序列：包括复制起始点(如大肠杆菌 *Ecol*1 复制起始点)、抗性基因(如氨苄抗性基因)及多克隆位点。

(2) 杆状病毒启动子：常用的有 P10 启动子和多角体启动子(*PPH*)。有些转移载体中可同时携带两个甚至多个杆状病毒启动子，以启动多种外源基因的同时表达。

(3) 杆状病毒复制非必需区：多采用多角体蛋白的侧翼序列，提供转移载体与杆状病毒胞内同源重组的位点。

(4) 昆虫细胞加尾信号。

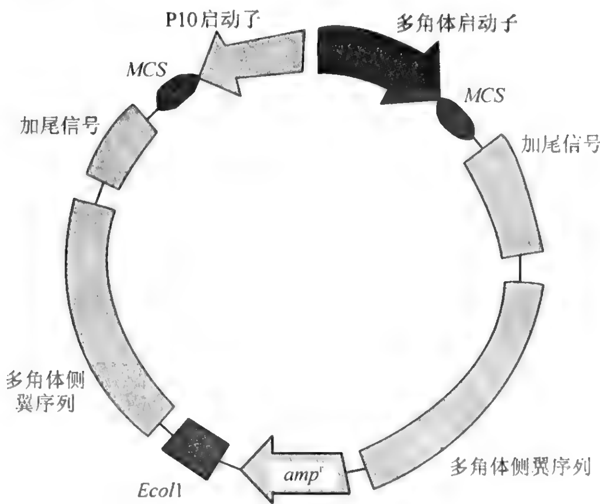


图 8-9　杆状病毒转移载体物理图谱

(三) 杆状病毒载体构建

建立杆状病毒载体必须将外源基因插入病毒基因组特定区域,通常是多角体编码区。因为杆状病毒基因组庞大,难以在体外拼接。首先必须将病毒 DNA 有关片段克隆于质粒 DNA 并做精细加工,切除多角体编码区,并保留其两端片段,接上多克隆位点,构建成转移载体质粒(transfer vector plasmid);随后,在体外拼接上外源基因,并利用共转染的方法,将野生型病毒基因和含有外源基因的重组转移载体质粒 DNA 同时导入细胞,由于两种基因都具有多角体编码区两端片段,二者在细胞内进行等位基因交换,外源基因插入到杆状病毒基因组,而将多角体蛋白基因灭活(图 8-10)。

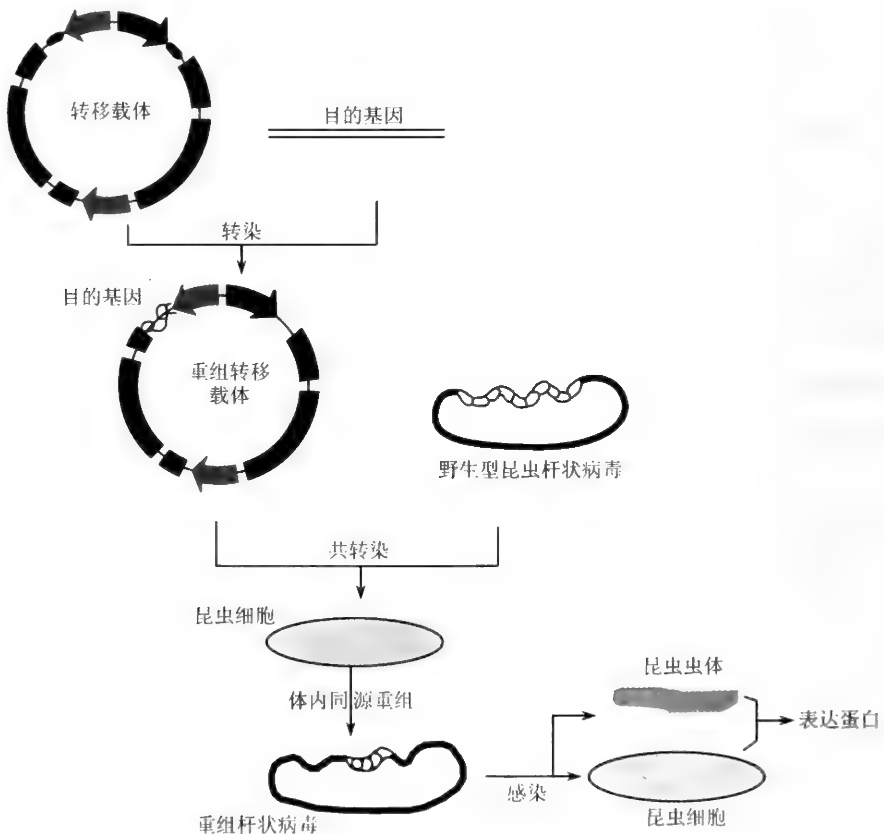


图 8-10 以杆状病毒为载体表达目的基因的流程图

二、在昆虫细胞中表达外源基因

杆状病毒载体系统的受体是昆虫细胞,目前这一表达系统已日益成熟,成为一种良好的高效表达系统。

在昆虫细胞中表达外源基因,首先必须构建一个合适转移载体(图 8-10)。

1. 转移载体质粒　此质粒中不含昆虫基因启动子，只是起转移作用，通过共转染与野生型病毒共同进入宿主，并在宿主内与野生病毒 DNA 发生同源交换，最终将外源基因插入野生型病毒基因组中。

2. 共转染　同时将病毒 DNA 和质粒 DNA 导入细胞的方法，称为共转染(co-transfection)。常用共转染的方法有磷酸钙沉淀法和脂质体技术、微量注射技术。其中磷酸钙沉淀法简便易行，转化效率较高，应用最广泛。

3. 重组病毒的筛选　由于重组病毒中外源基因插入多角体基因而使重组病毒表现为多角体阴性(Ocu^-)与野生病毒的多角体阳性(Ocu^+)区别明显。因而在培养的单层细胞上铺盖低融点琼脂糖，并在显微镜下挑选 Ocu^- 空斑，即可获得纯化的重组病毒。

基因产物的表达：以纯化重组病毒感染昆虫细胞或活体昆虫，即可在细胞培养上清中或细胞质内获得表达产物。

三、杆状病毒载体系统优势

杆状病毒载体系统比其他原核或真核表达系统具有明显优势：

1. 容量大　杆状病毒不但自身基因组大，其复制非必需区也较大，可以允许两种外源基因插入；

2. 具有强启动子　多角体蛋白启动子启动能力极强，其启动合成的多角体蛋白占细胞总蛋白的 50% 以上；

3. 有易识别标记　野生型病毒表达多角体蛋白(Ocu^+)，在显微镜下明显可见，重组病毒中外源基因的插入使多角体蛋白基因失活，不再产生多角体(Ocu^-)，这一标志非常明显，有利于阳性重组子的筛选；

4. 具有良好的后修饰作用，可使表达产物进行磷酸化、糖基化、切除信号肽等。

5. 安全　杆状病毒只感染昆虫细胞，对人畜无害，具有可靠的安全性。

（马春红　韩丽辉）

第九章　聚合酶链反应

Chapter 9. Polymerase Chain Reaction

聚合酶链反应(polymerase chain reaction,PCR)是一种在体外扩增特定基因或DNA序列的方法,又称基因体外扩增法,具有特异性、高效性和高度准确性三大特点。它为基因的分析与研究提供了一种强有力的手段,对现代分子生物学,乃至整个生命科学的研究与发展都有着深远的影响。PCR技术自1983年由美国PE-Cetus公司的Mullis等人发明至今,已被广泛应用于分子克隆、序列分析、基因突变、遗传病、传染病、性传播性疾病及法医判定和考古研究等领域,具有广阔的应用前景。

第一节　PCR的基本原理和影响因素

一、PCR的基本原理

PCR技术实际上是在模板DNA、引物和四种脱氧核糖核苷三磷酸存在的条件下依赖于DNA聚合酶的酶促合成反应。具有特异、高度敏感、快速、无放射性及操作简单等特点。特异性依赖于与靶序列两端互补的寡核苷酸引物。

PCR由变性—退火—延伸三个基本反应步骤构成:①模板DNA的变性:反应物加热至92～96℃,使模板DNA或经PCR扩增形成的DNA双螺旋解离,成为单链;②模板DNA与引物的退火(复性):模板DNA经加热变性成单链后,温度降低使引物与模板DNA单链的互补序列配对结合,此步温度变化范围比较大(具体温度从37～65℃都有),由引物与靶序列的同源程度和寡核苷酸的碱基组成决定;③引物的延伸:DNA模板-引物结合物在*Taq*DNA聚合酶的作用下,以dNTP为反应原料,以靶序列为模板,按碱基配对与半保留复制原理,合成一条新的与模板DNA链互补的半保留复制链,重复变性—退火—延伸三过程,就可获得更多的“半保留复制链”,而且这种新链又可成为下次循环的模板(图9-1)。每完成一个循环需2～4min,2～3h就能将待扩目的基因扩增放大几百万倍。

二、PCR反应的基本条件及其对PCR的影响

(一)模板

PCR以DNA或RNA为模板进行核酸的体外扩增。不过RNA的扩增需首先反转录成cDNA,然后才能进入PCR循环。标准分子生物学方法制备的样品DNA(RNA),其纯度足以用于PCR,且不需要另外的纯化步骤。模板若为哺乳动物基因组DNA,则每次PCR需模

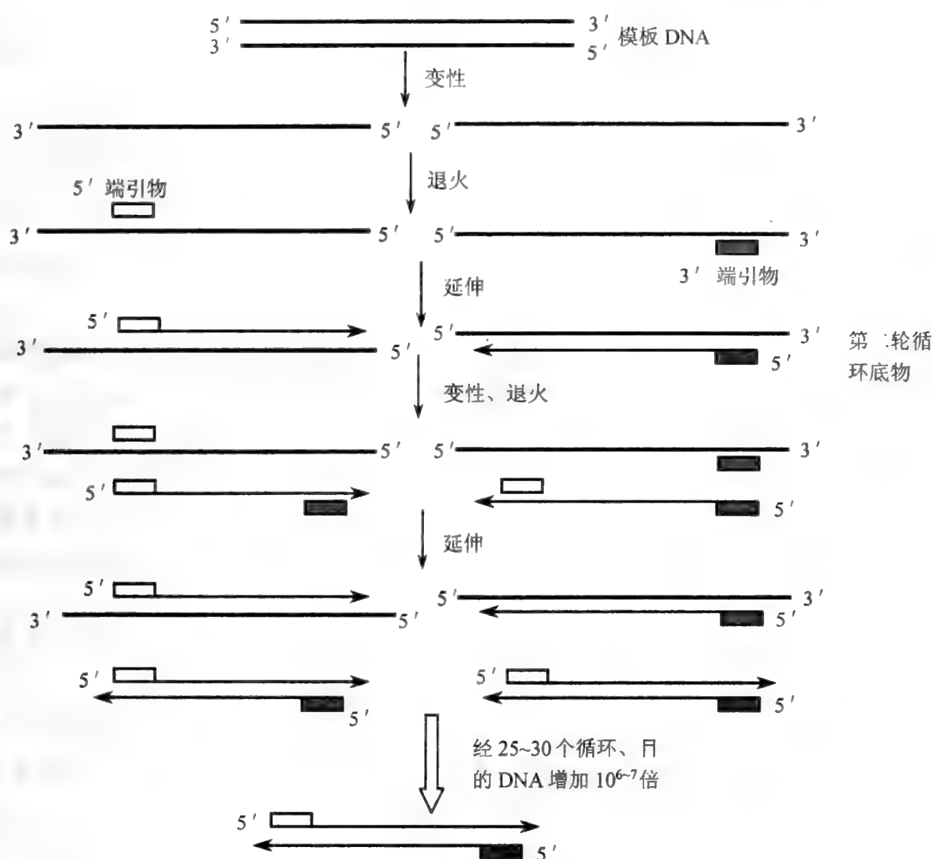


图 9-1　PCR 原理示意图

板量约为 0.1～1μg，若模板为细菌基因组或质粒 DNA，每次反应所需的量仅为 pg(10^{12}g)～ng(10^{9}g)即可满足要求。采集标本方法不当或靶 DNA 被丢失，都会导致出现假阴性，但如果标本中模板 DNA 浓度太高则会导致非特异扩增。

（二）引物

所谓引物，实际上就是两段与待扩增靶 DNA 序列互补的寡核苷酸片段，两引物间距离决定扩增片段的长度，两引物的 5′端决定扩增产物两端的位置。由此可见，PCR 结果的特异性取决于引物的特异性，扩增产物的大小也是由特异引物限定的。因此，引物的设计与合成对 PCR 的成功与否起着决定性的作用。

引物又分 5′端引物和 3′端引物，所谓 5′端引物是指与模板 5′端序列相同的寡核苷酸序列，又称为上游引物，3′端引物是指与模板 3′端序列互补的寡核苷酸序列，又称为下游引物。这两段序列一般为 15～30 个碱基。

引物设计一般遵循的原则：

1. 引物长度　根据统计学计算，长约 17 个碱基的寡核苷酸序列在人的基因组中可能出现的几率为 1 次。因此，一般情况下，引物长度不少于 16 个核苷酸，最高不超过 30 个核

苷酸，最佳长度为 20～24 个核苷酸。另外，可在引物 5′ 端添加不与模板互补的序列，如限制性酶切位点或启动子等，以完成基因克隆和其他特殊需要；引物 5′端生物素标记或荧光标记可用于微生物检测等各种目的。

2. (G+C)%含量　引物的组成应均匀，尽量避免含有相同的碱基多聚体。两个引物中(G+C)%含量应尽量相似，在已知扩增片段(G+C)%含量时宜接近于待扩增片段，一般以 40%～60%为佳。

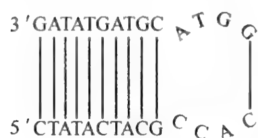


图 9-2　发夹结构示意

3. 引物内部应避免形成明显的次级结构，尤其是发夹结构(hairpin structures)。如图 9-2。

4. 两个引物之间不应发生互补，特别是在引物 3′端，即使无法避免，其 3′端互补碱基也不应大于 2 个碱基，否则易生成"引物二聚体"或"引物二倍体"(primer dimer)。所谓引物二聚体实质上是在 DNA 聚合酶作用下，一条引物在另一条引物序列上进行延伸所形成的与二条引物长度相近的双链 DNA 片段，是 PCR 常见的副产品，有时甚至成为主要产物。

另外，两条引物之间应避免有同源序列，尤其是连续 6 个以上相同碱基的寡核苷酸片段，否则两条引物会相互竞争模板的同一位点；同样，引物与待扩增靶 DNA 或样品 DNA 的其他序列也不能存在 6 个以上碱基的同源序列，否则，引物就会与其他位点结合，使特异扩增减少，非特异扩增增加。

5. 引物 3′端是引起延伸的起点，所以，引物 3′端 5～6 个碱基与靶 DNA 的配对要求必须精确和严格，这样才能保证 PCR 有效扩增。由于 AT、CG 引起的错配有一定规律，以 3′端"A"影响最大，因此尽量避免在引物 3′端第一位碱基是"A"。

6. 引物合成的质量是影响 PCR 成败的重要因素，人工合成的寡核苷酸引物最好经过色谱(层析)纯化或 PAGE 纯化，以除去未能合成至全长的短链等杂质。纯化引物在 25%乙腈溶液中 4 ℃保存可阻止微生物的生长；一般情况下，不用的引物应 -20℃保存，在液体中引物能保存 6 个月，冻干后可保存 1～2 年。

7. PCR 反应中引物的量也影响 PCR 扩增效果，如引物量太低，则 PCR 产物的量降低，甚至会出现假阴性。引物浓度过高不但会促进非特异产物合成，还会增加引物二聚体的形成。一般认为 PCR 反应中引物的终浓度为 0.1～0.5 μmol/L 为宜。

8. 引物 5′端可以修饰，如加酶切位点，用生物素、荧光物质、地高辛等标记，引入突变位点、引入启动子序列或蛋白质结合 DNA 序列。

(三) 缓冲液

PCR 反应缓冲液的目的是给 Taq DNA 聚合酶提供最适催化反应条件。目前最为常用的缓冲体系为 10～50 mmol/L Tris-HCl (pH8.3～8.8,20℃)，Tris 是一种双极性离子缓冲液，20℃时其 pK_a 值为 8.3，ΔpK_a 值为 -0.021/℃。因此，20 mmol/L Tris-HCl(pH 8.3, 20℃)，在实际 PCR 中，pH 变化于 6.8～7.8 之间。改变反应液的缓冲能力，如将 Tris 浓度加大到 50 mmol/L，pH8.9，有时会增加产量。

反应混合液中 50 mmol/L 以内的 KCl(pH8.9)有利于引物的退火，50 mmol/L NaCl 或 50 mmol/L 以上的 KCl 则抑制 *Taq* DNA 聚合酶的活性。反应中加入小牛血白蛋白(100

μg/ml)或明胶(0.01%)或 Tween 20(0.05%～0.1%)有助于酶的稳定性;反应中加入 5 mmol/L 的二硫苏糖醇(DTT)也有类似作用,尤其在扩增长片段(此时延伸时间长)时,加入这些酶保护剂对 PCR 反应是有利的。为了进一步提高扩增特异性,反应体系中可以加入二甲基亚砜(DMSO)或甲酰胺。

(四) Mg^{2+}

在 PCR 反应中,二价阳离子的使用是必需而严格的。研究表明,Mg^{2+} 离子优于 Mn^{2+} 离子,而 Ca^{2+} 离子无效。Mg^{2+} 浓度直接影响着酶的活性与准确性,从而影响反应的特异性和扩增片段的产率,在 PCR 反应中,Mg^{2+} 浓度一般在 1.5～2.0 mmol/L 范围内。另外 Mg^{2+} 浓度还影响引物的退火,模板与 PCR 产物的解链温度,产物的特异性,引物二聚体的形成等。Mg^{2+} 浓度过低时,酶活力显著降低;过高时,通常会导致非特异性扩增产物的累积。此外,PCR 混合物中的 DNA 模板、引物和 dNTP 的磷酸基团均可与 Mg^{2+} 结合,降低 Mg^{2+} 的实际浓度,而 *Taq* DNA 聚合酶需要的是游离 Mg^{2+}。所以无论是第一次靶序列与引物组合,还是 dNTP 或引物浓度发生改变时,均需优化 Mg^{2+} 浓度。

(五) 三磷酸脱氧核苷酸(dNTP)

四种脱氧核苷三磷酸(dATP、dCTP、dGTP、dTTP)是 DNA 合成的基本原料,四种 dNTP 的浓度应该相等,一般认为最适的 dNTP 浓度为 50～200 μmol/L。dNTP 含量太低,PCR 扩增产量太少、易出现假阴性;过高的 dNTP 浓度会导致聚合酶将其错误掺入。另外所需 dNTP 的量不但与靶序列的量有关,而且还与靶序列的大小有关,扩增大片段时应适当提高 dNTP 浓度。

(六) 耐热 DNA 聚合酶

1. PCR 之所以能得到广泛应用,主要是因为 *Taq* DNA 聚合酶取代了大肠杆菌 DNA 聚合酶 I 大片段。*Taq* DNA 聚合酶是从水生栖热菌 *Thermus aquaticus*(*Taq*)中分离出的热稳定性 DNA 聚合酶,使用 *Taq* DNA 聚合酶不仅简化了 PCR 程序,也极大地增加了 PCR 特异性及 PCR 扩增效率。该酶的最适温度很高(70～75 ℃),使引物在高温下进行退火和延伸,这样便增加了反应的总强度,并减少了与错配引物的延伸。在 PCR 反应中,每 100 μl 反应混合物中含 1～2.5 U *Taq* DNA 聚合酶为佳,酶的浓度太低,会使扩增产物产量降低,如果酶的浓度太高,则会出现非特异性扩增。*Taq* DNA 聚合酶保存不当而失活是 PCR 实验失败的常见原因。

2. T4DNA 聚合酶　在模板及引物存在的条件下,T4DNA 聚合酶催化沿 5′→3′方向 DNA 的合成。本酶还具有 3′→5′外切核酸酶的活性,但无 5′→3′外切核酸酶的活性,活性比 DNA 聚合酶 I 强。

3. T7DNA 聚合酶(基因 5 蛋白-硫氧环蛋白复合物)　T7 DNA 聚合酶由两种不同的亚基组成的:一种是 T7 噬菌体编码的基因 5 蛋白,另一种是大肠杆菌编码的硫氧环蛋白。前者同时具有 5′→3′聚合酶活性和 3′→5′外切核酸酶的活性,后者是一种辅助蛋白质,增强基因 5 蛋白质与引物模板的亲和性能,可使 DNA 聚合至数千个核苷酸长度。

4. $Vent_R$ DNA 聚合酶和 VentR(exo-)DNA 聚合酶　$Vent_R$ DNA 聚合酶和 VentR(exo-)DNA 聚合酶均纯化自重组大肠杆菌菌株,含有来自 archaea *Thermococcus litoralis* 菌的 Vent DNA 聚合酶基因,archaea *Thermococcus litoralis* 菌分离于海底火山口,能在高达 98℃ 的条件下生长。

$Vent_R$ DNA 聚合酶是一个高保真的耐热 DNA 聚合酶,它的保真性是 *Taq* DNA 聚合酶的 5～15 倍。该酶具有高保真性的部分原因是由于其自身具有 3′→5′核酸外切酶活性。在 95℃ 温育 1 小时后,该酶仍具有 90% 以上的聚合酶活性。

$Vent_R$(exo-)DNA 聚合酶是 $Vent_R$ DNA 聚合酶经基因工程改造,去除了 3′→5′核酸外切酶活性。该酶更适合于高温双脱氧测序反应和要求高产率的扩增反应,是应用于 DNA 测序的优选酶。

(七) 温度循环参数

1. 解链温度与时间　PCR 反应中模板 DNA 的变性十分重要,是 PCR 反应成败的关键因素之一。只有模板 DNA 和 PCR 产物双链完全解开,才能有效地和引物结合,这种结合是 PCR 扩增的基础。解链温度越高,时间越长变性就越充分,但温度过高、时间过长又会影响 *Taq* DNA 聚合酶的活性,所以通常选用解链温度为 94 ℃ 30 秒为宜。因模板 DNA 的链比较长,因此在 PCR 反应中第一个循环变性时间较长,多采用 94℃,5 分钟。

2. 复性(退火)温度与时间　复性温度是决定 PCR 反应特异性的重要因素。温度越低越好,但是容易出现引物与靶 DNA 的错配,增加非特异性结合,温度太高则不利于复性。复性温度与引物模板配对的 T_m 值有关,其最简单的计算方法是:$T_m = 4(G + C) + 2(A + C)$。在实践中由于 T_m 受到缓冲液中各种成分甚至引物、模板浓度不同的影响,因此任何计算得到的 T_m 值只能作为参考值。退火温度有一个温度范围(一般 2～5 ℃ 的跨度),一般从高到低直至低于 T_m 值 5 ℃ 的一系列反应中,可大致估计出最佳退火温度。复性时间并不是关键因素,但复性时间太长会增加非特异的复性。

3. 延伸温度与时间　引物延伸温度一般为 72 ℃。这个温度即考虑了 *Taq* DNA 聚合酶的活性,又考虑到引物和靶基因的结合。不合适的延伸温度不仅会影响扩增产物的特异性,也会影响其产量,72 ℃ 时,核苷酸的合成速度为 35～100 个核苷酸/s。72 ℃ 延伸 1 分钟对于长达 2 kb 的扩增片段是足够的。延伸时间过长会导致非特异性扩增带的出现。对很低浓度底物的扩增,延伸时间要长些。

4. 循环数　循环数决定着扩增的产量。在其他参数都已优化条件下,最适循环数取决于靶序列初始浓度。靶序列的初始浓度较低时,要增加循环次数。另外,酶活性不好或酶量不足时也要增加循环次数,以便达到有效的扩增量。

三、PCR 标本的制备

一般情况下,PCR 反应对 DNA 模板要求不严格。首先,反应混合物中存在一定量的蛋白质对实验过程影响不大,所以只要没有交叉污染,模板 DNA 制备可以不必像克隆、酶切、连接和标记等反应所用 DNA 制备那样严格。其次,模板用量很低,理论上 10^2～10^4 拷贝的

模板即可满足 PCR 反应的需要。但实际操作过程中，模板要求达到一定的水平，这样可以减少由于实验操作和试验精确度等方面原因引起的扩增失败。

四、PCR 实验中常见问题及对策

（一）假阴性

出现假阴性结果最常见的原因是：*Taq* DNA 聚合酶活力不够，或活性受到抑制；引物设计不合理；提取的模板质量或数量不过关以及 PCR 系统欠妥当；循环次数不够等。为了防止假阴性的出现，在选用 *Taq* DNA 聚合酶时，要注意用活力高、质量好的酶。尽管 *Taq* DNA 聚合酶对模板纯度要求不高，但不允许有破坏性有机试剂（如酚、氯仿等）的污染。PCR 扩增的先决条件及特异性是引物与靶 DNA 良好的互补，尤其是要绝对保证引物的 3′端与靶基因的互补。对变异较大的扩增对象，宜采用巢式（nested 或 double）PCR。在进行 PCR 操作时，应注意 Mg^{2+} 的浓度和各温度点的设置要合理。

（二）假阳性

因为 PCR 技术高度灵敏，极其微量的靶基因污染都会造成大量扩增，而造成结果判断上的失误，所以污染是 PCR 假阳性的主要根源。PCR 的污染主要是标本间的交叉污染和扩增子的污染。出现假阳性结果的另一种可能是样品中存在有靶基因的同源序列。为了避免因污染而造成的假阳性，PCR 操作时要隔离不同操作区、分装试剂、简化操作程序，使用一次性吸头。

第二节　常用的几种 PCR 技术

一、反 转 录 PCR

反转录 PCR（reverse transcription PCR，RT-PCR）即以 RNA 分子为模板的扩增技术，其过程包括，首先进行反转录将 mRNA 反转录为 cDNA，然后进行常规的 PCR 反应（图 9-3）。RT-PCR 主要用于克隆 cDNA、检测 RNA 病毒、合成 cDNA 探针及构建 RNA 高效转录系统。

（一）反转录酶

RT-PCR 的关键步骤是 RNA 反转录。必须注意用于反转录模板的 RNA 不能混有 DNA、蛋白质等，尤其不能混有 RNA 酶，所以要严格纯化 RNA。

RT-PCR 依赖于反转录酶将 mRNA 转录成 cDNA，所以选择合适的反转录酶对于 RNA 最大限度的转录为 cDNA 至关重要。常用的反转录酶有 AMV（鸟原粒细胞白血病病毒）、MMLV（莫洛尼鼠类白血病病毒）来源的反转录酶和来源于嗜热栖热菌 *Thermus thermophilus*、黄栖热菌 *Thermus flavus* 等嗜热微生物的热稳定性反转录酶。对于具有二级结构的 RNA，应选择最适反应温度较高的反转录酶，如 MMLV 反转录酶的突变株 *Rna*seH，

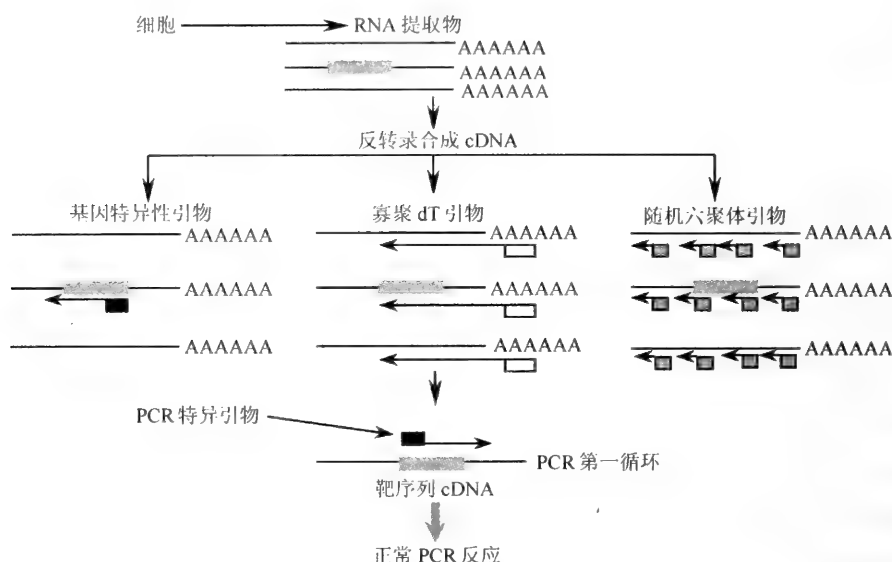


图 9-3 RT-PCR 反应原理

其最适反应温度为 50℃,此时可以破坏 RNA 的二级结构,能更有效地进行反转录。

(二) 反转录引物

1. 随机引物六聚体　当特定 mRNA 中含有反转录酶终止序列时,可以选用随机引物六聚体,使所有 RNA 均成为反转录的模板。

2. 寡聚(dT)　是一种对 mRNA 特异的引物。该引物可以与绝大多数真核生物 mRNA 配对结合,从而将其转录成 cDNA。

3. 特异引物　所谓特异引物,即能与目的 RNA 特异结合的寡核苷酸序列。PCR 反应一般需要一对引物,cDNA 第一条链可由 3′端引物(与 mRNA 3′端互补配对)起始转录,产生目的 RNA 特异的 cDNA 第一条链。

(三) 注意事项

1. 模板选择　并不是每种组织细胞均表达所有的基因,所以必须确定所选材料中含有所要扩增的模板基因。

2. 避免 RNA 酶(RNase)污染　提取或检测 RNA 时,所用器皿均需用 1‰焦碳酸二乙酯(DEPC)浸泡抑制 RNase 活性。

3. 引物　合成 cDNA 第一条链时,引物可用随机 6 聚核苷酸,或下游引物,或 poly-(dT),其中以 6 聚核苷酸引物效果最好。

二、定量 PCR

定量 PCR(quantitative PCR)的基本原理是,假定其反应产物的数量同反应混合物中起

始模板的 mRNA 或 DNA 的量成正比，因此通过琼脂糖电泳样品条带的比较，便可确定两种 PCR 产物之间的数量关系。定量 PCR 技术有广义和狭义之分。广义定量 PCR 技术是指以外参或内参为标准，通过对 PCR 终产物的分析或 PCR 过程的监测，进行 PCR 起始模板量的定量。

1．广义定量 PCR 常用技术

(1) 外参法＋终产物分析：所谓“外参法”是指样本与阳性参照在两个反应容器内反应。这种类型没有对样本进行质控监测，易出现假阴或假阳性结果，没有监测扩增效率，定量不准。

(2) 内参法＋终产物分析：所谓“内参法”是指样本与阳性参照在一个反应容器内反应，这种类型对样本进行质控监测，排除假阴性结果，但是定量不准。

(3) 内参法＋过程监测：由于样本与阳性参照在一个容器内反应，用同样的 *Taq* 酶和反应参与物，存在竞争性抑制，起始模板量浓度高的反应会抑制起始模板量浓度低的反应，所以定量不准。

2．狭义定量 PCR 技术　(严格意义的定量 PCR 技术)是指通过用外标法(如荧光杂交探针保证特异性)监测 PCR 过程(监测扩增效率)达到精确定量起始模板数的目的，同时以内对照有效排除假阴性结果(扩增效率为零)。

三、实时荧光定量 PCR

所谓实时荧光定量 PCR(real-time quantitative PCR)技术，是指在 PCR 反应体系中加入荧光基团，利用荧光信号积累实时监测整个 PCR 进程，最后通过校正曲线对未知模板进行定量分析的方法。实时荧光定量 PCR 技术于 1996 年由美国 Applied Biosystems 公司推出，由于该技术不仅实现了 PCR 从定性到定量的飞跃，而且与常规 PCR 相比，它具有特异性更强、有效解决了 PCR 污染问题、自动化程度高等特点，目前已得到广泛应用。

1．C_t 值的定义　在荧光定量 PCR 技术中，有一个很重要的概念——C_t 值。C 代表循环数(cycle)，t 代表阈值(threshold)，C_t 值的含义是：每个反应管内的荧光信号到达设定的阈值时所经历的循环数。

2．荧光阈值(threshold)的设定　PCR 反应的前 15 个循环的荧光信号作为荧光本底信号，荧光阈值的缺省设置是 3～15 个循环的荧光信号的标准偏差的 10 倍。

3．C_t 值与起始模板的关系　研究表明，每个模板的 C_t 值与该模板的起始拷贝数的对数成线性关系，起始拷贝数越多，C_t 值越小。利用已知起始拷贝数的标准品可做出校正曲线，其中横坐标代表起始拷贝数的对数，纵坐标代表 C_t 值。因此，只要获得未知样品的表 C_t 值，即可从校正曲线上计算出该样品的起始拷贝数。

4．PCR 过程的监测

监测 PCR 过程的方法有两种：荧光探针监测和荧光染料监测。荧光探针有三种即 *Taq*Man 荧光探针、杂交探针和分子信标探针，其中 *Taq*Man 荧光探针应用最广泛，本章主要介绍 *Taq*Man 荧光探针原理。

(1) *Taq*Man 荧光探针监测：PCR 扩增时在加入一对引物的同时加入一个特异性的荧

光探针，该探针为一寡核苷酸，两端分别标记一个报告荧光基团(reporter)和一个淬灭荧光基团(quencher)。探针完整时，报告基团发射的荧光信号被淬灭基团吸收；PCR 扩增时，*Taq* 酶的 5′→3′外切酶活性将探针酶切降解，使报告荧光基团和淬灭荧光基团分离，从而荧光监测系统可接收到荧光信号，即每扩增一条 DNA 链，就有一个荧光分子形成，实现了荧光信号的累积与 PCR 产物形成完全同步。如图 9-4 所示。

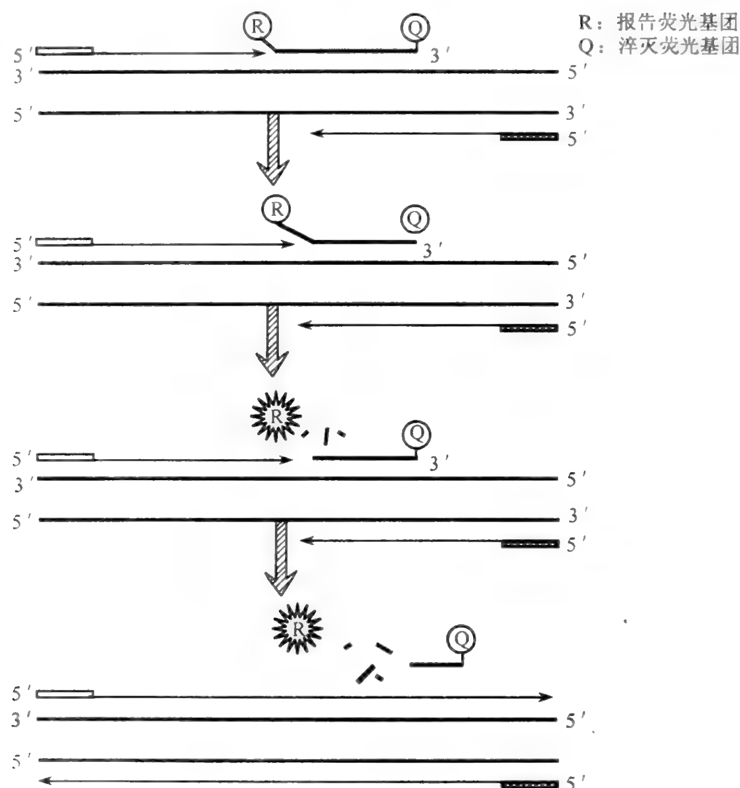


图 9-4 *Taq*Man 荧光探针原理示意图

(2) 分子信标探针检测：分子信标探针(molecular beacon)长约 25 个单核苷酸，空间结构呈茎环结构，结构可分环序列、茎杆部分和连接茎杆部位两端的荧光分子。环序列是与靶序列互补的探针，长约 18～20 个核苷酸；茎杆序列是与靶序列无关的链内互补序列，长约 5～7 个核苷酸；连接茎杆部位两端的荧光分子，一端为荧光分子，另一端为荧光淬灭分子，一般 5′端为荧光分子。

(3) 杂交探针检测：在普通 PCR 基础上增加两段荧光标记探针，两探针与靶 DNA 序列结合后相隔 1～2 个碱基。结合靶序列位于 5′端探针的 3′端有一荧光分子，位于 3′端探针的 5′端有一荧光分子。两荧光分子结合后发出荧光，游离时不发荧光。

(4) SYBR 绿色荧光染料监测：在 PCR 反应体系中，加入过量 SYBR 荧光染料，SYBR 荧光染料特异性地掺入 DNA 双链后，发射荧光信号，而不掺入 DNA 双链的 SYBR 染料分子不会发射任何荧光信号，从而保证荧光信号的增加与 PCR 产物的增加完全同步。

四、重组 PCR

重组 PCR(recombinant PCR,RP-CR)是指用 PCR 法在 DNA 片段上进行点突变,即扩增产物中含有与模板序列不同的碱基。由 PCR 扩增产生的核苷酸突变包括碱基替代、缺失或插入等。重组 PCR 需要两对引物,如图 9-5 分别设计"左方"、"右方"两对引物,分别为 a(5′端引物)、b(3′端引物)和 b′(5′端引物)、c(3′端引物),其中 b 和 b′含有互补的突变位点。先用两对引物对模板进行扩增,除去引物,将两种产物混合、变性、复性并延伸,然后再加入两端引物 a 和 c 进行常规扩增,最后得到中间部位特定位点发生突变的 DNA 片段。同样利用重组 PCR 可造成 DNA 片段的碱基插入或缺失,从而研究目的基因片段的功能。

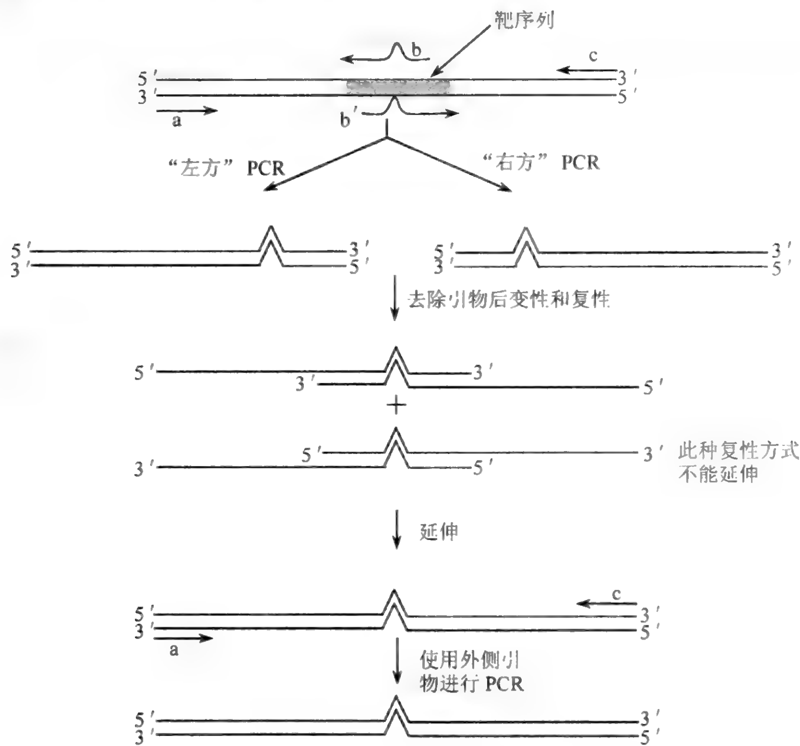


图 9-5 重组 PCR 导致特定碱基替换示意图

五、反向 PCR

反向 PCR(inverse PCR,IP-CR)是对一个已知的 DNA 片段核心序列两侧的未知序列进行扩增和研究的技术。选择已知序列内部没有切点的限制酶对此段 DNA 进行酶切,利用连接酶使之形成环状 DNA 分子,具已知序列设计合成合适的 3′ 端和 5′ 端引物扩增未知序列(图 9-6)。

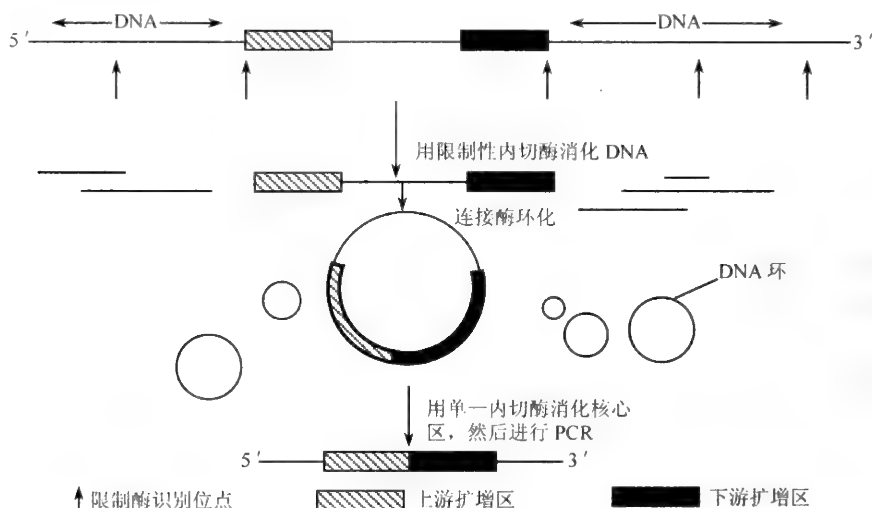


图 9-6　反向 PCR 的原理示意图

六、多重 PCR

多重 PCR(multiplex PCR)即在同一反应体系中加入多对引物,扩增同一模板的几个区域。如果基因的某一区段缺失,则相应的电泳图谱也缺失,多重 PCR 和 Southern blotting 一样可靠,但操作简单得多。多重 PCR 可同时检测多个突变位点或病原生物,有利于遗传病和感染性疾病的诊断。

七、长距离 PCR

一般情况下,正常 PCR 使用的 DNA 聚合酶只具有 5′→3′聚合活性,而无 3′→5′校正活性,这样就限制了 PCR 扩增片段的长度。长距离 PCR(long PCR)使用了两种 DNA 聚合酶,一种聚合酶无校正活性,另一种聚合酶具有 3′→5′方向校正活性,且其浓度较低,这样使 PCR 扩增片段长度大为增加。

八、不对称 PCR

不对称 PCR(asymmetric PCR)即在扩增体系中加入不同浓度的 3′端和 5′引物,而得到单链 DNA 产物。一般两条引物的浓度采用 50～100∶1 的比例。在最初 10～15 个循环中主要产物还是双链 DNA,但当低浓度引物被耗尽后,高浓度引物介导的 PCR 反应会产生大量的单链 DNA(ssDNA),纯化单链 DNA 即可进行其他试验。

第三节 PCR的应用

PCR技术自从建立以来,由于其敏感、高效、操作简便,在分子生物学研究和临床医学中得到广泛应用。

一、PCR在分子生物学中的应用

(一) 制备cDNA文库与筛选

用传统方法构建cDNA文库与筛选,既费钱费时,又需要较多的组织。但如果用PCR方法,只用很少组织甚至几个细胞就可以构建cDNA文库,并且大大提高了工作效率。

(二) DNA测序

PCR技术使DNA测序大为简化,因此,现在几乎都采用PCR法进行DNA序列测定。测序时,PCR产物既可以克隆到测序载体中,也可以不克隆直接测序。直接测序方便可靠,且可标准化和自动化,所以直接测序更普遍。

自动化测序仪器,大都使用荧光或放射性标记的核苷酸掺入法,然后进行电泳与自动化记录结果,自动测序中常用的反应是不对称PCR。

(三) 检测突变碱基

DNA碱基突变可引起肿瘤、遗传病、免疫性疾病等,因此,检测DNA突变分子对临床诊断与研究有重大意义。突变分子的检测有多种方法,如寡核苷酸探针(ASO)、限制性内切酶片段长度多态性(RFLP)、变性凝胶电泳(DGCE)及其他方法,但是以上各种方法,均不如聚合酶链反应-单链构象多态性(single strand conformation polymorphism analysis of polymerase chain reaction products,PCR-SSCP)分析技术简便、快速、经济。

PCR-SSCP技术的原理是:在非变性条件下,DNA单链在凝胶电泳中的迁移率与其相对分子质量和空间构象有关。DNA分子中碱基突变(即使只有一个碱基)可导致空间构象的改变,影响DNA的泳动速度,从而将变异的DNA与非变异DNA区分开来。

由于PCR-SSCP技术是检测单链DNA的,因此可应用不对称PCR,将PCR产物进行聚丙烯酰胺凝胶电泳,此时双链DNA泳动速度最快,其次是正常的单链DNA分子,最后是突变分子。

(四) 基因重组与融合

在分子生物学研究中常常需要将两个不同的基因融合在一起,通过PCR反应可以比较容易地实现这一目的。在两个PCR扩增体系中,两对引物分别由其中一对的5′末端引物和另一对的3′末端引物带上一段互补的序列。混合两种PCR扩增产物,在连接酶的作用下于互补端发生粘连,得到包含两个不同基因的DNA片段。

（五）用简并引物法扩增未知序列

简并引物是指按照氨基酸密码子简并性而合成的一套不同碱基序列。最常使用的情况有：当某一基因的序列不明，但其表达的蛋白质已纯化，并已对氨基末端的氨基酸测序，知道了几个氨基酸序列（通常为 5 个或以上氨基末端氨基酸序列）时；或者是为寻找某一基因家族未知成员时；或对某一基因家族分型鉴定时。使用简并引物的 PCR 反应，最适条件往往要凭经验确定，尤其是选定复性温度，以避免引物与模板之间发生错配，有的学者认为用热启动法能够有效地克服错配现象。

（六）其他

PCR 技术在其他很多方面都得到了应用，如基因定量、鉴定与调控蛋白结合的 DNA 序列；检测基因修饰，还可制备 DNA 探针、构建克隆或表达载体等。

二、PCR 技术在临床医学中的应用

（一）病原体诊断

利用 PCR 可以检测标本中的病毒、细菌、真菌、支原体，直至寄生虫等病原生物，标本可以是组织、细胞、血液、分泌物或排泄物等。对于某些病毒或细菌，如肝炎病毒等还可以分型。诊断的关键是要正确设计引物，通常应设置阳性对照与阴性对照，以防止出现假阴性或者假阳性结果。

如果模板 DNA 含量太少，可以进行巢式 PCR，提高 PCR 反应的敏感性、特异性和检测的准确性。

（二）遗传病基因的诊断

遗传病一旦发病，往往难以治愈，终生伴随，甚至夺去病人的生命，所以及早发现尤为重要。用 PCR 进行诊断，由于其成本低、快速和对样品质量和数量要求不高等特点，可在妊娠早期取得少量样品（如羊水）进行操作，发现异常胎儿可早期终止妊娠。

（三）组织器官移植的配型选择

器官移植尤其骨髓移植，对某些疾病的治疗具有独特的作用。同基因移植不发生免疫排斥，而异基因移植时，会发生移植排斥反应。避免移植排斥反应，需要找到与受者主要组织相容性抗原（HLA）相适应的个体。HLA 经典的配型方法是通过血清学或混合淋巴细胞培养方法分析 HLA 的基因表型，操作费时又不方便。使用 PCR 配型时，可使用有关 HLA 基因引物，对 HLA 进行扩增，并利用寡核苷酸探针进行杂交，可方便精确地选择出配型合适的供体。

（四）肿瘤的诊断、转移的确定

根据各种肿瘤细胞内基因突变的情况设计引物，进行 PCR，同时据正常基因设计引物

进行 PCR,比较二者 PCR 产物琼脂糖凝胶电泳结果,即可判断是否为肿瘤。若为血液标本、淋巴结或肿瘤临近组织即可活检,亦可有助于确定是否有肿瘤转移,由于 PCR 高度敏感,只需少量肿瘤细胞即可获得阳性结果,有利于早期诊断。

(五) 法医学

法医学应用 DNA 分析的目的包括个人识别和亲子鉴定。PCR 技术的应用和发展,有效地解决了法医学上取证遇到的困难。随着新的 PCR 技术的建立,必将促进法医学更快的发展。

(六) 其他领域

PCR 技术在动植物学研究领域、组织和群体生物学及古生物学等领域已得到广泛的应用。

(朱法良)

第十章 核酸杂交

Chapter10. Nucleic Acid Hybridization

核酸分子杂交技术 Nucleic Acid Hybridization 是现代分子生物学领域最常用的基本技术方法之一。其基本原理是具有一定同源性的两条核酸单链在一定的条件下(适宜的温度和离子强度),按碱基互补原则退火形成双链。

由于核酸分子杂交的高度特异性及检测方法的高度灵敏性,使核酸分子杂交技术在分子生物学领域中被广泛应用,如基因克隆的筛选和酶切图谱制作、基因组中特定基因序列的定量和定性检测、基因突变分析以及疾病的诊断等。核酸杂交的应用大大推进了分子生物学的迅猛发展。

第一节 核酸分子探针的标记

一、探针的种类及其选择

核酸分子杂交是利用已知探针检测待测核酸序列。其中待测核酸序列可以是克隆的基因片段,也可以是未克隆的基因组 DNA 或细胞总 RNA。核酸分子探针则是用放射性核素或其他标记物标记的、能与特定的靶分子发生特异性结合的 DNA 或 RNA 片段。

根据核酸分子探针的来源和性质可分为基因组 DNA 探针、cDNA 探针、RNA 探针及人工合成的寡核苷酸探针等几类。根据目的要求的不同,可以采用不同类型的核酸探针。值得注意的是,并不是任意一段核酸片段均可作为探针。探针选择的最基本原则是探针必须具有高度的特异性,另外还要考虑来源是否方便等其他因素。

(一) 基因组 DNA 探针

以基因组 DNA 作为探针在目前应用最为广泛。基因组是某一生物细胞全套基因的组合,几乎所有的基因片段都可被克隆到质粒或噬菌体中,这为获得大量高纯度的基因组 DNA 探针提供了非常方便的来源。近年来发展起来的多聚酶链反应(PCR)为基因组 DNA 探针的制备提供了另一方便的来源。但在选择此类探针时,要特别注意真核生物基因组中存在的高度重复序列(如人类基因组中的 *Alu* 序列)。要尽可能使用基因的编码序列(外显子)作为探针,而避免使用内含子及其他非编码序列,否则探针可能因高度重复序列存在引起非特异性杂交而出现假阳性结果。

（二）cDNA 探针

cDNA 也可作为探针。cDNA 是以 mRNA 为模板，在反转录酶的作用下形成的互补 DNA(complementary DNA，简称 cDNA)。由于 cDNA 不存在内含子及其他高度重复序列，因此是一种较为理想的核酸探针。但因 cDNA 探针不易获得，从而限制了它的广泛应用。另外，以 cDNA 为探针还必须注意 poly(dT)产生的非特异性杂交问题。

（三）RNA 探针

mRNA 作为核酸分子杂交的探针具有以下优点：

1. RNA/RNA 和 RNA/DNA 杂交体的稳定性较 DNA/DNA 杂交体的稳定性高，因此杂交反应可在更为严格的条件下进行(杂交温度可提高 10℃左右)，因而杂交的特异性更高。

2. 由于单链 RNA 分子不存在互补双链的竞争性结合，其与待测的核酸序列杂交的效率更高。

3. RNA 不存在高度重复序列，因此非特异性杂交较少。

4. 杂交后可用 RNase 将未杂交的探针分子消化掉，从而使本底降低。

但是，大多数 mRNA 存在一多聚腺苷酸尾，有时也会影响其杂交的特异性。此缺点可通过在杂交液中加入 poly(A)将待测核酸序列中可能存在的 poly(dT)或 poly(U)封闭而加以克服。另外，RNA 极易被环境中大量存在的核酸酶降解，因此不易操作，这也是限制其广泛应用的重要原因之一。

（四）寡核苷酸探针

近年来，随着 DNA 合成仪的问世及使用的逐步推广，人工合成的寡核苷酸探针的应用越来越广泛。其优点是：

1. 可根据需要合成相应序列，避免了天然核酸探针中存在的高度重复序列带来的不良影响。

2. 特别适合于基因点突变分析，因为大多数寡核苷酸探针长度只有 15～30bp，其中只有一个碱基不配对也会显著影响其 T_m 值(melting temperature，解链温度，即 DNA 溶液一半的双链 DNA 已解离为单链的温度)。

3. 由于合成的寡核苷酸探针链短，序列的复杂性低，相对分子质量小，因此杂交所需时间也较短。

但是，需要注意的是，短寡核苷酸探针所带的标记物较少，特别是非放射性标记时，其灵敏度较低。

通常设计寡核苷酸探针的方法有两种：

1. 根据已知的核酸序列设计寡核苷酸探针　随着人类基因组计划的实施和完成，目前已有大量的基因被克隆并进行了序列测定。因此，对相当多的基因而言，可通过查阅基因库(gene bank)来合成一段寡核苷酸序列。其原则是：

(1) 一般要求探针长度为 10～50bp。过短会导致特异性降低，过长则不仅合成较困难，

而且杂交时间也会延长。

(2) G:C 含量应为 40%～60%，超出此范围会增加非特异性杂交。

(3) 不应有大于 4bp 的碱基反向互补配对，否则探针内部会形成“发夹样”结构。

(4) 避免有同一碱基的连续出现(如—G-G-G-G—)。

(5) 最好在计算机中与已知的各种基因序列进行同源性比较，如寡核苷酸探针序列与非靶基因序列有 70%以上的同源性或连续有 8 个以上的碱基序列相同，则最好不用，重新设计。

2. 从蛋白质多肽序列推测寡核苷酸序列　对于一些尚未克隆的基因，可根据其蛋白质氨基酸序列合成相应的寡核苷酸序列作为探针。由于氨基酸密码的简并性，有些氨基酸可由多种密码子编码，这使我们根据蛋白质的氨基酸序列来推测其基因的碱基顺序的工作变得甚为困难。要完全准确地将其基因顺序推测出来几乎是不可能的。但有以下几个原则可供参考：

(1) 尽量选择那些只含有一种密码子编码的氨基酸(如色氨酸和蛋氨酸)的多肽作为合成寡核苷酸的参照。

(2) 尽管某些氨基酸可由多种密码子编码，但其使用频率是不一样的，因此应优先考虑使用频率最高的密码子。

(3) 真核基因中极少有 CG 序列，因而，密码子内部凡是存在 CG 顺序的密码使用率极低。另外，当两个相邻的氨基酸的最高频率密码子相连导致 CG 顺序出现时，应将其中一个氨基酸密码换为次高频密码子。通常将前一个密码子 NNC 换成 NNT，甘氨酸、苏氨酸除外。

(4) 参考同一基因家族的基因顺序，对于密码子的正确选择也是有所帮助的。

(5) 采用上述原则合成的寡核苷酸探针一般错配率较低，选择合适的杂交条件(如适当降低杂交温度等)，可基本满足克隆筛选等工作的需要。为了尽可能减少假阴性结果，也可合成各种可能的寡核苷酸探针分别或混合进行探测。当混合进行杂交探测时，由于各探针的(G+C)%含量不同，因此需分别在不同的温度下进行探测或采用适当的方法使其 T_m 值基本趋于相同。

二、各种标记物及其选择

一种理想的探针标记物，应具备以下几种特性：①高度灵敏性；②标记物与核酸探针分子的结合，应绝对不能影响其碱基配对特异性；③不影响探针分子的主要理化特性，特别是杂交特异性和稳定性，杂交体的解链温度(T_m)应无较大改变；④不影响 Ag-Ab 或酶和底物的反应；⑤标记及检测方法简单、保存时间长；⑥对环境无污染，对人体无损伤，价格低廉。

可用于核酸探针的标记物主要有两大类，即放射性标记物和非放射性标记物。

(一) 放射性标记物

放射性核素的灵敏度极高，可检测到 0.01pg 甚至更微量的核酸分子。其最大的优点是放射性核素与相应的元素具有相同的化学性质。由于放射性核素与相应元素的差别只在中子数目上，质子和电子数完全一致，而元素的化学性质是由其核外电子决定的，因此放射性核素对各种酶促反应无任何影响，也不会影响碱基配对的特异性与稳定性及杂交性质。另外，放射性核素的检测具有极高的特异性，少数假阳性结果的出现，极少是由于放射性核素

引起的，而主要是由杂交过程本身导致的。因此若严格按规程操作（主要是预杂交和洗膜），则假阳性率极低。放射性核素的主要缺点是易造成放射性污染。另外，多数放射性核素的半衰期都较短，因此必须随用随标，标记后立即使用，不能长期存放。

常用于标记核酸探针的放射性核素有^{32}P、^{35}S、^{3}H，下面简要介绍这几种放射性核素：

1. ^{32}P　能产生穿透力极强的β射线，因此放射自显影所需时间短，灵敏度高，分辨率高，是最常用的放射性标记物。^{32}P常以^{32}P-dNTP或^{32}P-NTP的形式提供，标记部位有α位、γ位标记。前者常用于切口平移法、随机引物法等标记法，后者用于多核苷酸激酶末端转移法的标记。^{32}P的缺点是半衰期短（约14d），射线散射严重，导致X线片上带形轮廓不清，有时影响结果的分析，特别是当要求高分辨率时（如DNA序列测定和细胞原位杂交时）。

2. ^{35}S　S原子可以取代磷酸分子上的一个氧原子，从而形成^{35}S标记的核苷酸分子。分子结构的这种改变对大多数核酸修饰酶（DNA和RNA聚合酶、激酶和磷酸酶等）的活性没有太大的影响，因此可替代^{32}P标记的核苷酸用于探针标记。另外，因^{35}S射线的散射作用较弱，在X线胶片上成影分辨率较高，且半衰期长（约87d），故常用于核酸序列测定及细胞原位杂交。但因^{35}S释放β粒子能量低，粒子穿透力极弱，检测的灵敏度较^{32}P稍弱。

3. ^{3}H　^{3}H释放的β粒子能量极低，散射极少，因此在感光乳胶上的成影分辨率高，本底较低，最适用于细胞原位杂交。且其半衰期长，标记的探针可存放较长时间反复使用。但因放射自显影所需时间较长，其应用受到限制。

（二）非放射性标记物

近几年来，随着非放射性标记物和标记技术的发展，非放射性标记物标记探针的敏感度接近放射性核素标记探针的敏感度，并且具有操作简便，获得结果快，无放射性污染，稳定性较好，可以长时间存放的优点，目前广泛应用于基础和临床研究，有逐渐取代放射性核素标记探针的趋势。根据检测方法的不同，非放射性标记物可分为以下几类：

1. 半抗原　目前常用的非放射性标记物有生物素和地高辛，它们都是半抗原，可以利用这些半抗原的抗体进行免疫检测。

（1）生物素标记探针是通过一个长臂（核苷酸与生物素之间的长链）使生物素连接到核苷酸嘧啶环的C-5或嘌呤的N-6位上，制成生物素化的核苷酸，然后以一定方法使标记的核苷酸掺入到探针中去，而不影响杂交过程中的碱基配对，杂交后加入酶标的抗生物素蛋白，酶分解底物出现有色反应而检测。即生物素化核苷酸→掺入到DNA中去制成探针→杂交→加入酶标抗生物素蛋白→加酶底物→显色。生物素与核苷酸间的长臂的作用是避免生物素分子与核酸分子结合后因空间障碍影响与抗生物素蛋白的结合，灵敏度检测结果表明臂长与灵敏度有一定关系，其中以Bio-16或11-dUTP标记的DNA探针的敏感度最高，可达1～2pg。利用生物素标记的探针优点是无放射性，操作简便，但缺点是敏感性低，原位杂交时因组织细胞存在有内源性生物素，易出现假阳性结果。

（2）地高辛，是一种类固醇半抗原，像生物素一样，可通过一长臂人工结合于dUTP上，然后以一定方法使标记的dUTP掺入到DNA中去，得到地高辛标记的核酸探针，探针与样品中核酸杂交后加入酶-抗地高辛抗体复合物，最后通过酶使底物显色而检测。其优点是无放射性，操作简便安全，灵敏度接近同位素标记（达0.1pg），特异性较高，因组织细胞内无内

源性地高辛,所以原位杂交时,不易出现假阳性。

2. 配体　生物素还是一种抗生物素蛋白(卵白素,avidin)和链霉素抗生物素蛋白(streptavidin)的配体,可以用亲和法进行检测。

3. 荧光素　如 FITC、罗丹明类,可以被紫外线激发出荧光进行观察,主要适用于细胞原位杂交。

4. 光密度或电子密度标记物　如金、银等,适用于细胞原位杂交,可以在光镜下或电镜下进行观察。

5. 还有一些标记物可与另一些物质反应而产生化学发光现象,可以像放射性核素一样直接对 X 线胶片进行曝光。如 Promega 公司生产的 Lightsmith^{TM}II Luminescence Engineering System 及 Amersham 公司的 ECL 等。这类标记物可能是今后研究主流。

三、探针标记方法

探针的标记方法可分为体内标记法和体外标记法。体内标记法是将放射性标记的化合物作为代谢底物加到活细胞中,经细胞的代谢处理而将生物大分子加以标记。例如将^{3}H-胸苷加入到活细胞培养基中,可标记 DNA;而^{3}H-尿苷则可标记 RNA。体内标记技术对于追踪某些分子在细胞中的代谢途径是极其有效的手段。但由于此方法受到多种因素的限制,使标记化合物的比放射活性不高。另外,不同生物细胞的代谢途径不一样,要选择不同的标记底物才能得到有效的标记。因而,在分子生物学领域中普遍使用的核酸分子探针几乎都是用体外标记法标记的。体外标记法有两种:

(一) 化学标记法

化学标记法是利用标记物分子上的活性基团与探针分子上的基团(磷酸基团)发生的化学反应而将标记物直接结合到探针分子上,如光敏生物素的标记。光敏生物素是指将生物素与另一类化学性质较活泼的基团(叠氮基或补骨脂)相连。光敏生物素与核酸探针混合后,在强可见光(汞灯或钨灯)照射下(10~20min),光敏生物素中的活泼基团与核酸共价相连,成为生物素标记的核酸探针。

(二)酶促标记法

酶促标记法是将标记物预先标记到核苷酸分子上,然后利用酶促方法将核苷酸分子掺入到探针分子中,或将核苷酸分子上的标记基团交换到探针分子上。此法适用于所有放射性核素对核苷酸的标记,部分非放射性标记的核苷酸分子也可用此法进行标记,如生物素、地高辛标记的核苷酸。下面介绍几种常用的酶促标记法:

1. 缺口翻译法或切口平移法　切口平移法是实验室中最常用的一种脱氧核糖核酸探针标记法。它利用大肠杆菌 DNA 聚合酶 I 的多种酶促活性将标记的 dNTP 掺入到新形成的 DNA 链中,从而合成高比活性的均匀标记的 DNA 探针。线状、超螺旋及带缺口的环状双链 DNA 均可作为切口平移法的模板。

其基本原理是以极微量的 DNA 酶Ⅰ(DNase Ⅰ)在 Mg^{2+} 的存在下，在待标记的双股 DNA 的每一条链上产生若干单链缺口，再利用大肠杆菌 DNA 聚合酶Ⅰ的 5′→3′核酸外切酶活性及 5′→3′聚合酶活性，在缺口处 5′末端每切除一个核苷酸，则在 3′末端添加一个核苷酸，以修补缺口，随着缺口在 DNA 上的移动，标记的核苷酸就掺入到新合成的 DNA 链中(图 10-1)。

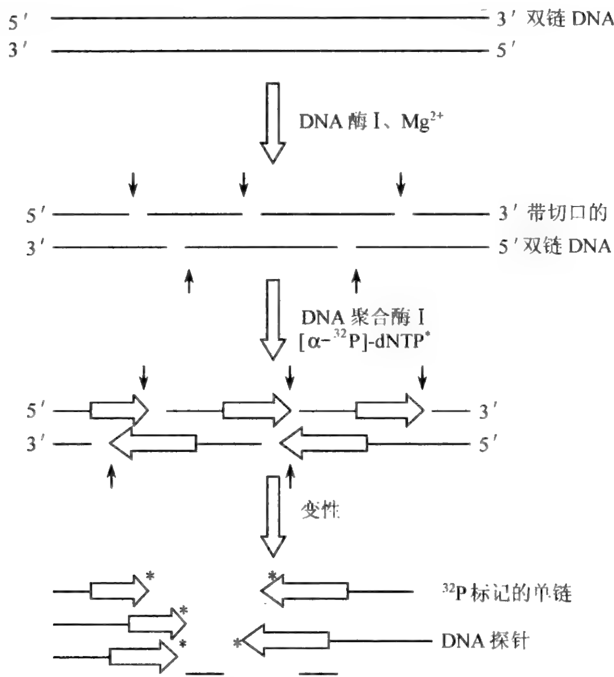


图 10-1　切口平移法原理示意图

注：⇦* ^{32}P 标记的 DNA

此方法的特点是快速、简便、标记探针均匀，特异性高，最常用于标记双股 DNA 探针。单链 DNA 或 RNA 及小片段(100～200bp)双链 DNA 不能用此法标记。

需注意的是，DNA 多聚酶必须是大肠杆菌 DNA 多聚酶全酶，此酶的 Klenow 大片段不具有 5′→3′外切酶的活性，所以不能替代。再者，DNase Ⅰ的浓度一定要适当。DNase Ⅰ浓度过大，将导致 DNA 链上形成的切口过多，从而使探针长度过短，影响杂交反应的效率；DNase Ⅰ的浓度过低，则不足以形成足量的切口，将会使标记效率下降。理想的结果应使 30%～60% 的核苷酸掺入到 DNA 中，最后形成的单链 DNA 探针的理想长度为 400～800bp。

2. 随机引物法　随机引物法是近年来发展起来的一种较理想的核酸探针标记方法，目前大有取代缺口平移法而作为实验室中 DNA 探针标记常规方法的趋势。其基本原理是利用随机寡核苷酸引物和大肠杆菌 DNA 聚合酶Ⅰ的 Klenow 片段来标记 DNA。所谓随机引物是含有各种可能排列顺序的寡核苷酸片段的混合物，它可以与任意核苷酸序列杂交，起到聚合酶反应引物的作用。随机引物可用 DNase Ⅰ降解小牛胸腺 DNA 获得(6～8 个核苷酸)，或用人工合成方法获得(6

个核苷酸)。当待标记的DNA探针变性后与随机引物一起杂交时,随机引物按碱基互补配对原则各自与模板上的不同区段结合,Klenow片段即从引物3′-OH端开始合成互补DNA链,反应中若加入修饰的dNTP即可获得标记的DNA探针(图10-2)。

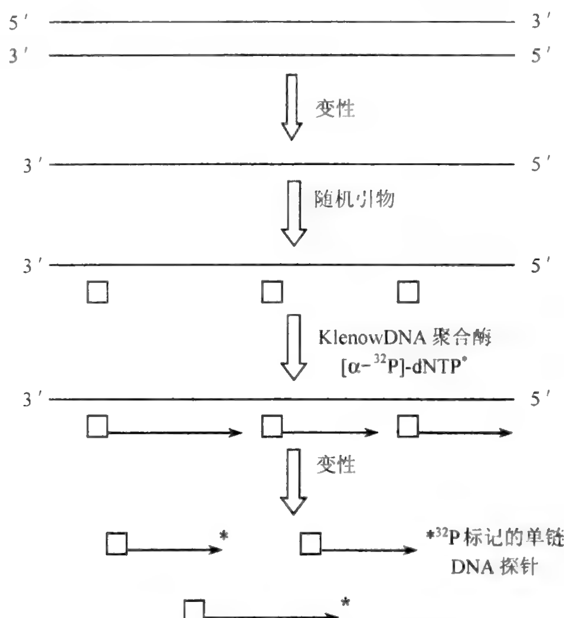


图10-2 随机引物标记法原理示意图

注: □ 随机引物 ⟶* ^{32}P标记的DNA

此方法的特点是:①不但可以用于双链DNA标记,而且可用于单链DNA或RNA探针的标记;②操作方便,避免了用DNase Ⅰ处理时,酶浓度掌握不当所带来的一系列问题;③标记效率高,3h内可使40%~60%以上甚至90%的修饰DNA掺入到探针DNA中去。④因核苷酸掺入率极高,故探针可不经Sephadex-50纯化即可用于杂交,还可直接在低熔点琼脂糖溶液中进行标记。⑤由于反应条件控制在pH 6.6,可抑制KlenowDNA聚合酶3′→5′外切酶活性,延长反应时间也不会造成DNA的降解。必要时可将反应时间延长到12~16h,使更多的核苷酸掺入到DNA中,从而提高标记活性。

3. 寡核苷酸探针的末端标记　目前,人工合成的寡核苷酸片段作为分子杂交的探针日益受到人们重视。利用寡核苷酸探针可以检测到靶基因上单个核苷酸的点突变。用于寡核苷酸探针标记的方法有很多,在此仅介绍常用的三种:

(1) 末端脱氧核苷酰转移酶法(简称末端转移酶法):末端转移酶可催化[α-^{32}P]-dNTP在单链DNA 3′末端的多聚化(详见第四章)。

此方法标记活性高,适用于克隆筛选点突变分析及细胞原位杂交等,缺点是不能控制增加的核苷酸的数量,使得到的标记产物长短不一,但可加入标记的[α-^{32}P]-ddNTP而控制。

(2)T4多核苷酸激酶法:先用碱性磷酸酶将5′端P去掉或加入过量ADP接受5′-P形成ATP,而暴露5′-OH。再利用T4多核苷酸激酶的作用催化ATP分子上的γ-磷酸基因转移到DNA或RNA分子的5′-OH上,采用γ-^{32}P-ATP为底物,即可将DNA样品5′末端标记(详见第四章)。

此方法的缺点是在每个探针的5′末端多加了一个磷酸，理论上，这会影响其与DNA的杂交。再者，因生物素等标记物是连在碱基上，而不是磷酸基团，因此T4多核苷酸激酶法不可用于生物素等的末端标记。

（3）Klenow DNA多聚酶标记法：对于带黏性末端的双链寡核苷酸可利用Klenow DNA多聚酶的填充反应进行末端标记。而对于单链寡核苷酸，则可预先合成一小段与此探针互补的寡核苷酸作为引物，然后利用Klenow DNA多聚酶的链延伸反应获得标记的寡核苷酸探针。此法的特点是模板DNA与标记的DNA探针长度不相同。如果需要，可采用电泳方法将它们分离开来。

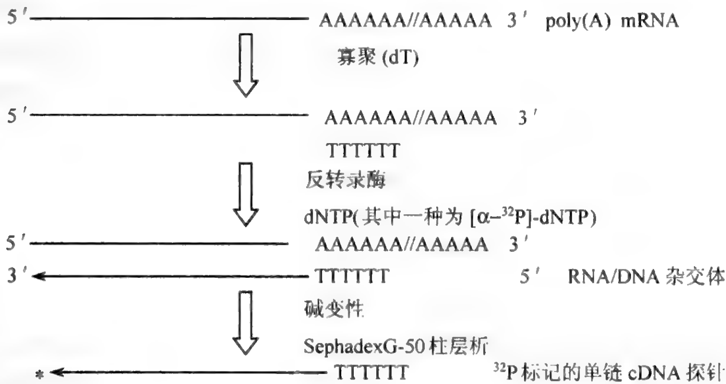


图10-3 ^{32}P标记的单链cDNA探针的制备

注：*← ^{32}P标记的DNA

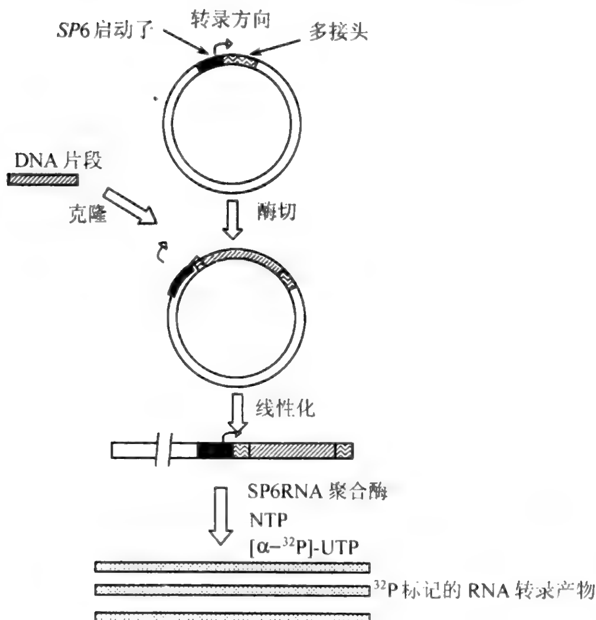


图10-4 采用*SP6*DNA载体体外转录体系合成^{32}P标记的RNA探针

4. cDNA的标记 常以mRNA为模板,以寡聚dT为引物,也可用特异的寡核苷酸引物或随机寡核苷酸做引物在反转录酶的作用下,进行反转录,反转录得到的产物RNA/DNA杂交双链经碱变性后,RNA单链可被迅速降解成小片段,经SephadexG-50柱层析,即可获得单链cDNA探针(图10-3)。

5. RNA探针的标记 常用体外转录的方法制备RNA探针。利用*SP*6RNA聚合酶体系,首先将感兴趣的基因序列克隆到*SP*6表达载体的*SP*6启动子下游的多克隆位点下,用适当的限制性内切酶在插入序列的下游将质粒线性化。*SP*6噬菌体的RNA聚合酶对*SP*6启动子序列具有高度的亲和性,从而启动其下游序列的转录(图10-4)。因此,如果在4种NTP(其中一种用^{32}P标记)的存在下,利用*SP*6RNA聚合酶的Run-off转录活性,即可获得高放射活性的RNA探针,最后DNA模板可用无RNase污染的DNase处理而清除掉。

第二节 核酸分子杂交的技术

一、核酸分子杂交的类型

核酸分子杂交可分为液相杂交和固相杂交两种。液相杂交是在液体中进行的杂交方法,杂交速度快,常与核酸电镜技术结合在一起,研究不同DNA的同源性及mRNA与染色体DNA间的关系等。液相杂交最大缺点是不能区分探针DNA与目的DNA,给常规应用带来困难。固相杂交是在一定支持物上进行的杂交反应,其优点是检测方便,因而应用广泛。它包括膜固相印迹杂交和原位杂交两种。

(一)膜固相印迹杂交

膜固相印迹杂交是指将待测核酸序列结合到一定的固相支持物上,然后与存在于液相中的标记核酸探针进行杂交的过程,是目前最常用的一种核酸分子杂交方法。选择良好的固相支持物与有效的转移方法是此项技术成败的关键。下面介绍固相支持物的选择及常用的膜固相印迹杂交类型。

1. 固相支持物的选择

(1)硝酸纤维素滤膜:硝酸纤维素滤膜具有较强的吸附单链DNA和RNA的能力,特别是在高盐浓度下,其结合能力可达80~100$\mu g/cm^2$。吸附的单链DNA或RNA经真空烘烤后,依靠疏水性相互作用而结合在硝酸纤维素膜上。另外,硝酸纤维素膜还具有杂交信号本底较低的优点,因此被广泛应用于Southern、Northern、斑点印迹及克隆筛选中。硝酸纤维素膜非特异性吸附蛋白质的作用较弱,特别适合于那些涉及蛋白质作用(如抗体和酶等)的非放射性探针标记的杂交体系。

硝酸纤维素滤膜是应用最广泛的一种固相支持物,但并不十分理想。因为硝酸纤维素膜是依靠疏水性相互作用结合DNA的,这种结合并不十分牢固,随着杂交及洗膜的进程,DNA会慢慢脱离硝酸纤维素膜,特别是高温情况下,从而使杂交效率下降。因此,不太适宜于同一膜上重复进行杂交。再者,硝酸纤维素膜质地较脆,特别是经烘烤后,操作不方便,须特别小心。硝酸纤维素膜与核酸的结合有赖于高盐浓度(>10×SSC),在低盐浓度时,结合

DNA 的效果不佳,因此不适宜于电转移印迹法。另外,硝酸纤维素膜对于相对分子质量小的 DNA 片段(特别是小于 200bp 的 DNA 片段)结合能力不强。因此,现在更多的人倾向于使用尼龙膜。

(2)尼龙膜:尼龙膜是目前较理想的一种核酸固相支持物,它有多种类型,除网眼大小不一样外,有的尼龙膜未经特殊处理,有些则是经过了正电荷基团的修饰,这种修饰的尼龙膜结合核酸的能力更强。

尼龙膜结合单链及双链 DNA 和 RNA 的能力较硝酸纤维素膜更强,可达 350~500μg/cm^2。而且经烘烤或紫外线照射后,核酸可牢固地结合在尼龙膜上,这种结合较硝酸纤维素膜强,特别是用短波紫外线照射后,核酸中的部分嘧啶碱基可与膜上的带正电荷的氨基相互交联,从而使结合更为牢固。碱处理也可使 DNA 牢固地结合在尼龙膜上,因此使 DNA 的变性、吸印和固定可以一步完成。有报道微波处理也可使 DNA 结合在尼龙膜上,此方法特别适合于菌落原位吸印法,使细菌的裂解、DNA 变性与固定可一步完成(注意:硝酸纤维素膜不能用微波处理,否则会引起燃烧)。另外,它的韧性较强,操作较方便;对于相对分子质量小的核酸片段亦有较强的结合能力,甚至短至 10bp 的核酸片段也能结合;在低离子强度条件下也可较好地结合 DNA,因此比较适合于电转印迹法;再者,尼龙膜可重复用于杂交,一次杂交后,探针分子可经碱变性而被洗脱下来,从而可用于与第二个探针进行杂交。其缺点是杂交本底较高,可以用加大预杂交液中的非特异性封闭试剂的方法克服。

2. 固相印迹杂交的类型

(1) 斑点及狭缝印迹法(dot or slot blot):将 RNA 和 DNA 变性后直接点于硝酸纤维素膜上,用于基因组中特定基因及其表达的定性及半定量研究,称为斑点印迹。它是实验室常用技术之一。与 Southern 及 Northern 印迹法相比,其优点是简单、迅速,可在同一张膜上同时进行多个样品的检测;对于核酸粗提样品的检测效果也较好。缺点是不能鉴定所测基因的相对分子质量,不能确定外源 DNA 在细胞染色体中的整合情况,而且特异性不高,有一定比例的假阳性。

(2) Southern 印迹(Southern blot):Southern 印迹是指将电泳分离的 DNA 片段从凝胶中转移到一定的固相支持物上进行杂交的过程。该转印技术是 1975 年 Southern 建立的,故称为 Southern 印迹法。

DNA 分子经限制性内切酶酶切后,再经琼脂糖凝胶电泳将所得 DNA 片段按相对分子质量大小分离成不同区带,然后将含 DNA 片段的琼脂糖凝胶变性,并将其中的单链 DNA 片段转移到硝酸纤维素膜或其他固相支持物上,而各 DNA 片段的相对位置保持不变。这种膜即可用于下一步的杂交反应。Southern 转印过程可利用毛细管虹吸作用由转移缓冲液带动核酸分子转移到固相支持物上;也可利用电场作用进行电转;还可利用真空抽滤作用进行真空转移。

利用 Southern 印迹法可进行克隆基因的酶切图谱分析、基因组基因的定性及定量分析、基因突变分析及限制性片段长度多态分析(RFLP),外源基因的整合等。

(3) Northern 印迹(Northern blot):Northern 印迹是指将 RNA 变性及电泳分离后转移到固相支持物上进行杂交的过程,转移后的支持物可用于杂交反应以鉴定其中特定 mRNA 分子的量及大小。此转印技术是 1977 年由 G.R.Stark 建立的。

Northern 印迹的基本原理及过程与 Southern 印迹相同，前面提到的毛细管虹吸法、电转法及真空转移法也同样适用于 RNA 的转移。只是 RNA 变性方法与 DNA 不同，因为碱会导致 RNA 的水解，故不能用碱变性 RNA。RNA 变性常用甲醛或聚乙二醛或甲基氢氧化汞。RNA 变性后有利于在转印过程中与硝酸纤维素膜结合，它同样可在高盐中进行转印，但在烘烤前与膜结合并不牢固，所以在转印后不能用低盐缓冲液洗膜，否则 RNA 会被洗脱。在凝胶中不能加 EB，因为它会影响 RNA 与硝酸纤维素膜的结合，为了测定片段大小，可在同一凝胶上加标记物一同电泳，之后将标记物胶切下、染色，照相，含样品的凝胶进行 Northern 转印。

（二）原位杂交

原位杂交属于固相分子杂交的范畴，它是用标记的 DNA 或 RNA 探针，在原位检测组织或细胞内特定核酸序列的方法。包括组织细胞原位杂交、菌落原位杂交。

1. 组织细胞原位杂交　组织细胞原位杂交是在细胞保持基本形态的情况下，探针进入细胞内与 RNA 或 DNA 杂交。此杂交过程主要是在载玻片上进行的。根据其所用探针及所要检测的核酸不同，可分为 DNA-DNA、RNA-DNA 及 RNA-RNA 杂交。由于原位杂交能在成分复杂的组织中进行单一细胞的研究而不受同一组织中其他成分的影响，因此对于那些数量少且散在于其他组织中的细胞内 DNA 或 RNA 的研究更为方便；同时由于原位杂交不需要从组织中提取核酸，对于组织中含量极低的靶序列有极高的敏感性，并可完整地保持组织与细胞的形态，更能准确地反映出组织细胞间的相互关系及功能状态。

2. 菌落原位杂交　菌落原位杂交是将待检测的菌落转移到硝酸纤维素膜，然后再利用探针检测细菌菌落或噬菌斑中特异的 DNA 序列或基因，这在筛选重组 DNA 文库时尤为有用。其基本过程是先将要筛选的细菌菌落从琼脂平板上原封不动地转移到硝酸纤维素膜上，注意转移菌落时在琼脂平板和硝酸纤维素膜上的相应位置都要做标记，便于以后辨认原始菌落。膜上的菌经原位裂解、变性、固定、烤膜后，DNA 就固定在膜上，经杂交后确定含特异性基因的菌落。

二、核酸分子杂交的基本过程

核酸分子杂交实质上是双链 DNA 的变性和具有同源序列的两条单链的复性过程，其基本过程包括变性、预杂交、杂交、洗膜及检测。

（一）变性

维系 DNA 螺旋的力主要是氢键和疏水性相互作用。DNA 的变性是指用加热或强碱使双股 DNA 之间的氢键打开形成单链的过程。此时，DNA 的二级结构发生破坏，但未破坏其一级结构。热变性通常采用 95～100 ℃ 加热 5～10min；碱变性采用 0.5mol/L NaOH 作用 30～60min。

（二）预杂交

预杂交的目的是用非特异性 DNA 分子（鲑精 DNA 或小牛胸腺 DNA）及其他高分子化合物（Denharht 溶液）将待测核酸分子中的非特异性位点封闭，否则这些非特异性位点会在杂交时与探针非特异性结合。

（三）杂交

指待测 DNA 样品中若存在与加入的 DNA 探针同源的互补顺序，在一定的条件下，即可退火而形成异源 DNA 双链，即 DNA 杂交体。其实质上是 DNA 的复性过程。复性是指变性的两条单链 DNA 按碱基互补配对原则，在适当条件下重新缔合形成双链的过程。

复性并不是变性反应的简单逆反应过程。变性可以在一个极短的时间内迅速完成，而复性则需相对较长时间才能完成。如果使热变性的 DNA 溶液迅速冷却，则只能形成一些不规则的碱基对，而不会恢复 DNA 双链结构。而将此变性溶液在低于 T_m 值 25℃ 的温度下维持相当长的时间（24h 左右），则可使之回复至天然的双链结构状态。

复性时的第一步是两条 DNA 单链随机碰撞形成局部双链的过程，这种随机碰撞形成的局部双链是暂时的，如果局部双链周围的碱基不配对，则重新解离。一旦找到正确的互补区，首先形成的局部双链就形成核心，核心两侧的碱基顺序迅速配对，形成完整的双链分子。

杂交体系的建立应考虑以下因素：

1. 离子强度　通常采用 5× 或 6× SSC（1× SSC 为 0.15mol/L NaCl，0.015mol/L 枸橼酸钠），离子强度过高或过低可能影响杂交的特异性。

2. DNA 浓度　DNA 浓度愈高，复性速度愈快。为保证足够的 DNA 浓度，除应加入足够的 DNA 量以外，还应尽量减少杂交体积，一般每平方厘米滤膜面积加 50～100μl 杂交液为宜。

3. DNA 探针的长度　探针片段越大，其扩散的速度越慢，因此复性的速度越慢。

4. 杂交温度　选择适当的杂交和洗膜温度是核酸分子杂交成败最关键的因素之一。通常杂交反应在低于 T_m 值 15～25℃ 温度下进行。T_m 值除受（G+C）含量影响外，还与杂交体中离子的强度，探针的复杂性（长度），是否含有甲酰胺等多种因素有关。如果一长度为 500bp 的探针，（G+C）含量为 50%，杂交体系中含 5× SSC，不含甲酰胺，则 T_m 值为 93.4℃。那么其杂交温度（T*hyb*）应为 93.4－25＝68.4℃；若含有 50% 甲酰胺，则 T_m 值为 68.4℃。那么其杂交温度（T*hyb*）应为 68.4－25＝43.4℃。

一般情况下，在不含变性剂甲酰胺时，大多数杂交反应在 68℃ 进行，当含有 50% 甲酰胺时，在 42℃ 进行。

5. 提高杂交效率　在杂交液中加入硫酸葡聚糖，可促进链间的缔合。同时，由于本身的微粒结构可吸附探针分子，从而使接触面积增加，使杂交效率提高 10～100 倍；缺点是使杂交背景加深。也有人使用聚乙二醇（PEG，相对分子质量为 6000～8000）和聚丙烯酸（相对分子质量为 90 000），此两者的优点是黏稠度较硫酸葡聚糖低。

（四）洗膜

洗膜是在一定离子强度和适当温度下，洗去非特异性吸附在支持物上的 DNA 探针的

过程。洗膜温度的高低直接影响杂交的背景和敏感度,温度过高可能使特异性结合的探针脱落,而洗膜温度过低,则不能将非特异性吸附的探针洗掉。一般最终洗膜浓度应在低于 T_m 值 5~12℃下进行。例如,探针长 500bp,(G+C)含量为 42%,离子强度为 0.1×SSC,不含甲酰胺,T_m 值为 67℃,则洗膜温度为 55~62℃,对多数杂交体系而言,最典型的终洗膜浓度为 0.1×SSC,洗膜浓度为 55℃。

(五)杂交信号的检测

杂交信号的检测方法因探针上的标记物不同而异。通常分为以下两类:

1. 放射性核素探针的检测　采用放射自显影的方法检测,即利用放射性核素产生的射线(β或γ)使 X 线片感光而检测。其方法是将杂交膜用保鲜膜包好,在暗室中,与 X 线片压在一起,置于暗盒中,于 -70℃冰箱中曝光适当时间后,在暗室中取出 X 线片,经显影、定影后即可观察结果。

2. 非放射性核素探针的检测

(1)地高辛标记探针的检测:地高辛是一种半抗原,可通过抗原-抗体反应和酶显色体系偶联而检测。例如,杂交后,地高辛标记的探针即特异性结合在硝酸纤维膜上,检测时加入地高辛抗体-碱性磷酸酶复合物,抗体与地高辛抗原结合,当加入碱性磷酸酶的底物(NBT 和 X-磷酸盐)时,酶则分解底物产生蓝紫色不溶性反应物。

(2)生物素标记探针的检测:生物素(biotin)是抗生物素蛋白(avidin)的配体。抗生物素蛋白是一种糖蛋白,其分子中有 4 个与生物素结合的位点,与生物素的亲和力极高,生物素探针常通过亲和法进行检测。酶常用辣根过氧化物酶,其底物为二氨基联苯胺(DAB),在 H_2O_2 存在的情况下,辣根过氧化物酶分解 DAB 产生红棕色沉淀物。

(王晓燕)

第十一章 DNA 测 序

Chapter11：DNA sequencing

DNA 测序(DNA sequencing)，就是确定组成 DNA(脱氧核糖核酸)的四种化学碱基的确切顺序(即核酸一级结构的测定)，是现代分子生物学中的一项重要技术。常规测序方法的基本过程包括：第一，把待测序的 DNA 分子进行处理，得到只差 1 个核苷酸的一系列逐步缩短的 DNA 分子混合物；第二，通过凝胶电泳把这些 DNA 分子分离开来，形成阶梯状排列的条带，然后逐个读出 DNA 的碱基序列。

目前主要应用两种快速序列测定技术：1977 年由 F. Sanger 等提出的双脱氧链终止法和 1977 年由 Maxam 与 Gilbert 提出的化学降解法，其中双脱氧链终止法是目前应用最多的技术。近几年，由于 DNA 测序策略的日趋成熟，以及自动测序仪的改进，序列测定速度和准确率大为提高。

第一节 DNA 序列测定的策略

DNA 序列测定简便快速，但由于受到电泳凝胶分辨率的限制，一般每次测序只能测出 500～600 个核苷酸，因此对更长 DNA 分子的测定必须考虑到策略问题，以确保 DNA 序列分析简捷、快速、准确。

一、从遗传图谱、物理图谱到基因组序列图谱

以 DNA 测序为基础的基因组学研究主要包括遗传图谱、物理图谱的绘制和基因组序列的测定，以及解读、注释分析基因组序列等四个方面的内容。遗传图谱，即通过亲本的杂交，分析后代的基因或其他特异分子标记物间重组率，并用重组率来表示两个基因之间距离的线性连锁图谱。物理图谱是以大片段的基因组 DNA 分子为基本单位，以连续的重叠群为基本框架，通过遗传标记将重叠群或基因组 DNA 分子排列于染色体上，描绘 DNA 上可以识别的标记的位置和相互之间的距离(以碱基对的数目为衡量单位)，这些可以识别的标记包括限制性内切酶的酶切位点、基因等。对于人类基因组来说，最粗的物理图谱是染色体的条带染色模式，最精细的图谱是测出 DNA 的完整碱基序列。基因组序列是分辨率为单个碱基的物理图，也是最终基因图的主要结构基础。

在基因组测序中，存在着两种不同的战略，即全基因组随机测序战略和以物理图为基础的以大片段克隆为单位的定向测序战略。两者的最大区别在于是否依赖基因组作图。把每条染色体分成平均长度为 400 kb 的片段，每段克隆到一个酵母菌人工合成染色体(yeast

artificial chromosome,YAC)上,所有 YAC 克隆都按照其在染色体上的实际位置进行排序,可得到一个能够覆盖整个染色体的 YAC 文库。把每一个 YAC 克隆携带的染色体片段经部分酶切形成一系列有重叠区域的 40 kb 左右的片段克隆到黏粒上,得到黏粒文库。每个黏粒上的染色体片段再经酶切形成 4 kb 左右的片段克隆到测序专用的质粒载体上。测序质粒上携带的 4 kb 片段就可以用现在常规测序的方法进行测序了。把所有质粒克隆的 DNA 片段序列读出,再按照各个片段在染色体上的实际位置进行排列,最后就可以得到染色体的全部核苷酸碱基对序列。

二、鸟　枪　法

全基因组随机测序战略主要采用鸟枪法(shotgun),也俗称"霰弹法",是随机先将整个基因组打碎成小片段进行测序,最终利用计算机根据序列之间的重叠关系进行排序和组装,并确定它们在基因组中的正确位置。

采用超声波处理或酶解的方法,将待测的 DNA 片段随机地切成大小不同的小片段并亚克隆。分别测定核苷酸序列后,通过计算机处理各片段核苷酸序列的资料,最终根据重叠的序列连成待测 DNA 的全长序列。全基因组鸟枪法测序的主要步骤是:第一,建立高度随机、插入片段大小为 2 kb 左右的基因组文库。克隆数要达到一定数量,即经末端测序的克隆片段的碱基总数应达到基因组 5 倍以上。第二,高效、大规模的末端测序。第三,序列集合。第四,填补缺口。有两种待填补的缺口,一是没有相应模板 DNA 的物理缺口,二是有模板 DNA 但未测序的序列缺口。

"鸟枪法"优点是速度快,好像树林里停了一大群鸟,很多人乱枪射击,在很短的时间内,就可以将林子中的大部分鸟打中,而且简单易行,成本较低,可以在较短的时间内通过集中机器和人力的方法获得大量的基因片段。但是用它来测序,最终排序结果的拼接组装比较困难,尤其在部分重复序列较高的地方难度较大。此外,有许多序列片段难以定位在确切的染色体上,成为游离片段;同时又会有许多地方由于没有足够的覆盖率而形成空缺。这些缺陷最终导致整个基因图会留下大量的空洞(gap),也影响其准确度。

三、引物延伸法

引物延伸法(primer walking)是利用通用引物确定待测 DNA 片段两端的序列以后,根据测出片段的一端序列,设计新的引物再向前测序,以此方式不断向前测序,直到获得待测片段的全长序列。

四、套式缺失法

套式缺失法(即限制性酶切-亚克隆法)是采用核酸外切酶从待测 DNA 片段的一端逐个降解核苷酸,通过选择不同的降解时间,可以获得一套长短不同的单向缺失的突变株。突变株亚克隆后分别进行序列测定,然后将这一整套单向缺失突变株的核苷酸序列拼接起来,即

可获得待测 DNA 片段的序列。

第二节 常用的 DNA 测序方法

Sanger 双脱氧链终止法和 Maxam-Gilbert 化学修饰法是目前公认的两种最常用、最有效的 DNA 序列手工分析法。但目前利用现代精密仪器和机器人技术实现了 DNA 测序的高度自动化,市场上已经有各种型号的 DNA 自动测序仪可供选购。

一、Sanger 双脱氧链终止法

英国剑桥大学分子生物学实验室的 F. Sanger 等人于 1977 年发明了利用 DNA 聚合酶的双脱氧链终止原理测定核苷酸序列的方法。由于这种方法要求使用一种单链 DNA 模板和一种适当的 DNA 引物,因此有时也称为引物合成法。它利用了 DNA 聚合酶所具有的两种酶催化反应特性:第一,DNA 聚合酶能够利用单链 DNA 作模板,合成出准确的 DNA 互补链;第二,DNA 聚合酶能够利用 2′、3′-双脱氧核苷三磷酸作底物,将其掺入到寡核苷酸链的 3′末端,从而终止 DNA 链的延伸。

(一)Sanger 双脱氧链终止法原理

通常 DNA 复制所需要的组分包括:DNA 聚合酶、单链 DNA 模板、带有 3′-羟基末端的单链寡核苷酸引物、4 种 dNTP(dATP、dGTP、dTTP 和 dCTP)。以 DNA 单链为模板,在特定条件下,用特异的引物在 DNA 聚合酶的作用下,根据碱基互补配对的原则,不断将 4 种脱氧核糖核苷酸(dNTP)加到引物的 3′-羟基末端并使引物链得到延伸,合成出新的互补 DNA 链。这种链的延伸是通过引物的 3′-羟基和脱氧核糖核苷酸底物的 5′-磷酸基团形成磷酸二酯键来完成的。如果这种反应体系中加入双脱氧核糖核苷酸(ddNTP),这种 2′、3′-ddNTP 的 5′-磷酸基团是正常的,而 3′位置缺少羟基,因此在 DNA 聚合酶的作用下,仍然可以通过 5′-磷酸基团与引物链的 3′-羟基反应掺入到引物链中,但是由于 ddNTP 没有 3′-羟基,不能继续与下一个 5′-磷酸基团形成磷酸二酯键而导致引物链延伸的终止。例如,存在 ddCTP、dCTP 和三种其他 dNTP(其中一种为 α-^{32}P 标记)的情况下,将引物、模板和 DNA 聚合酶一起保温,即可形成一种全部具有相同 5′引物端和以 ddC 残基为 3′端结尾的一系列长短不一的片段混合物。这样,在测序反应体系中,DNA 引物链不断合成与偶然终止,产生一系列的长短不等的核苷酸链。然后将这些测序反应产物经变性聚丙烯酰胺凝胶电泳,分离制得的放射性自显影区带图谱将为新合成的不同长度 DNA 链中 C 的分布提供准确信息,从而将全部 C 的位置确定下来。采用类似的方法,在 ddATP、ddGTP 和 ddTTP 存在的条件下,可同时制得分别以 ddA、ddG 和 ddT 残基为 3′端结尾的三组长短不一的一系列核苷酸链片段,其长度取决于以起始 DNA 合成引物末端到出现过早链终止位置之间的距离,结果将产生 4 类寡核苷酸,它们分别终止于模板链的每一个 A、C、G 或 T 的位置上。将制得的四组混合物全部平行地分别点加在变性聚丙烯酰胺凝胶电泳板上进行电泳,每组制品中的各个组分将按其链长的不同得到分离,从而制得相应的放射性自显影图谱,根据自显影图谱、

从4组泳道上按终止反应时所形成不同长度DNA的区带位置依次读出从3′到5′的碱基序列(图11-1)。该过程可由计算机读写,也可人工分析。

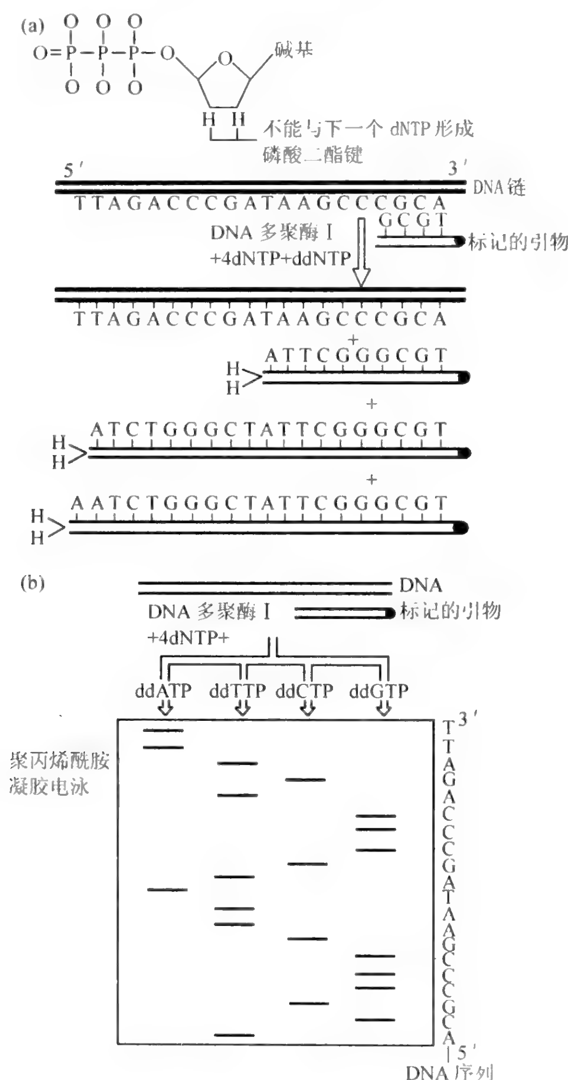


图11-1 双脱氧链末端终止法测序基本原理示意图

(a) 双脱氧核苷三磷酸的结构及碱基配对原则

(b) 测序反应过程及结果判定

(二)双脱氧链终止法测定步骤

以M13噬菌体为载体的双脱氧链终止法的实施步骤为例:

1. M13噬菌体复制型双链DNA(RFDNA)的制备 为了将待测双链DNA进行克隆,必须制备M13 mp18或M13 mp19复制型(闭环双链)DNA

2. 重组噬菌体制备 采用多克隆位点上的限制性内切酶切割待测 DNA,并以相同内切酶切割 M13 噬菌体 RF DNA,然后将待测外源 DNA 片段与 M13 复制型载体 DNA 连接、转化、筛选出阳性重组噬菌体。

3. 单链噬菌体 DNA(模板 DNA)的制备 上述平板的每个白色噬菌斑是由一种重组 M13 DNA 分子转染细菌后产生的,如果取一个白色噬菌斑进行培养便可得到一种单链形式的 DNA。

4. 引物制备 利用 DNA 合成仪合成一段与待测 DNA 片段 3′端互补的寡核苷酸引物。可以购买通用引物或实验室合成 15～26 bp 长度的引物。

5. 微量滴板 将待测单链 DNA 模板与寡核苷酸引物混合,引物与 DNA 单链模板的 3′端进行杂交。将上述杂交混合液分为四份,每份混合液中均加入等摩尔量的四种 dNTP,其中一种用放射性标记(通常用 α-^{32}P-dATP),此外还需分别加入一种双脱氧核糖核酸(ddATP,ddGTP,ddCTP 或 ddTTP),四管中的 ddNTP 均不同。加入 DNA 聚合酶反应数分钟,然后在适当温度条件下延伸互补链,就可得到四组分别以 A、G、C、T 终止的长短不一的放射性标记的互补链混合物,其范围从几个碱基对至 200 bp 左右。

6. 变性聚丙烯酰胺凝胶电泳 为得到测序结果,胶的质量和电泳条件与测序反应本身同样重要,没有一种实验方案适合于所有测序目的,所以测序胶的长度以及聚丙烯酰胺浓度应根据不同的实验而定。将四组含有长短不等的反应混合物在同一聚丙烯酰胺凝胶板上进行电泳,即可将只相差一个核苷酸的寡核苷酸链分离开。

7. 电泳后将凝胶干燥,放射自显影 将电泳凝胶进行放射自显影,即可得到放射自显图谱,通过该图谱即可直接读出核苷酸的排列顺序。

(三) 双脱氧链终止法的特点

在进行 DNA 顺序测定前,需要用不同的限制酶消化待测 DNA,使其降解成小片段,分别克隆到 pUC18、pUC19、M13mp18、M13mp19 等载体中,测定各小片段的顺序,由于不同限制酶产生的片段之间有交错重叠顺序,根据片段间和末端重叠序列,用计算机软件排列出各片段的位置,进而排出 DNA 的全序列。

对于双链 DNA 模板,在测序反应过程中只加入一条引物,这条引物只与双链 DNA 模板任意一条链配对进行测序反应,与引物不配对的另一条 DNA 单链在反应中不会干扰序列反应的测定。

在大量 DNA 测序时(如人类基因组计划),为节省时间和材料,可在同一凝胶的四条泳道上电泳几种不同的测序反应物。不同的反应经使用不同的引物开始而不是放射性标记物。电泳后,将测序胶转印到膜上,并用放射性标记的探针与其中的一种引物互补杂交,因而在与一种模板相应的四条泳道上产生信号。读序后,洗掉探针并进行下一个引物-探针杂交反应。

二、Maxam-Gilbert 化学法测定 DNA 序列

DNA 的化学修饰测序法是于 1977 年由美国哈佛大学的 A. Maxam 和 W. Gilbert 发明

的,所以又叫作 Maxam-Gilbert DNA 序列分析法。

(一) 原理

在 DNA 化学测序法中,与碱基发生专一性反应的化学试剂在一种或两种特定核苷酸位置上随机断裂 3′端或 5′端标记的 DNA 链,用化学试剂处理末端带有放射性标记的 DNA 片段,造成碱基的特异性切割,由此产生一组具有各种不同长度的 4 套寡聚脱氧核糖核苷酸 DNA 片段混合物,经凝胶电泳分离和放射自显影之后,便可根据 X 线底片上所显现的相应谱带,直接读出待测 DNA 链的核苷酸序列,可得到如图 11-2 所示的 4 种 DNA 阶梯。

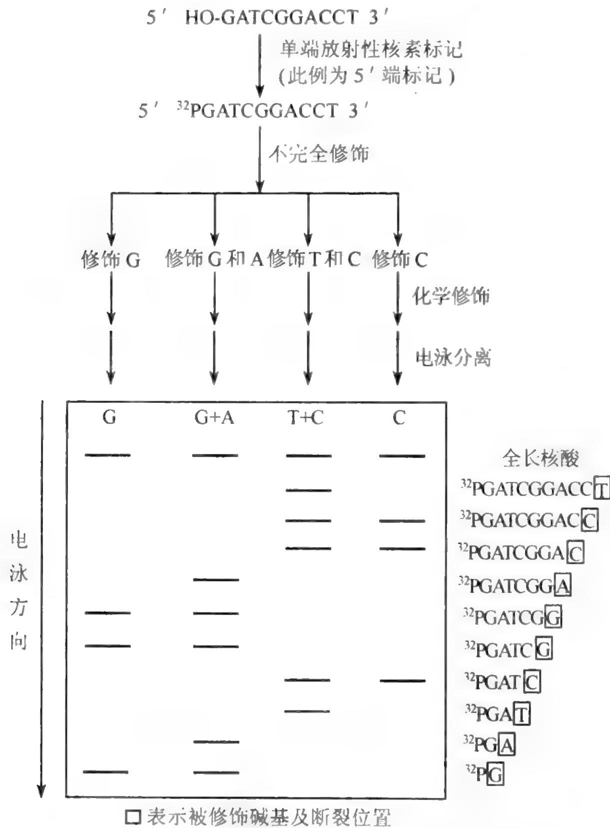


图 11-2 化学裂解法测定 DNA 的核苷酸序列

(二) 步骤

把待测序的 DNA 分子变成单链分子,其 5′端用 ^{32}P 进行放射性标记。然后把这些单链 DNA 分子分成 4 份,每份用一种化学试剂(肼、硫酸二甲酯或甲酸)处理 DNA 片段,每种试剂可使 DNA 分子在一种碱基 5′端的磷酸二酯键处发生断裂。把反应条件控制好,使每个 DNA 分子只发生一次断裂,就得到 4 种反应产物,每种由一种碱基处发生断裂形成的 DNA 片段组成。把这 4 种反应产物用聚丙烯酰胺凝胶电泳进行分离,电泳完成后,用 X 线胶片

进行曝光,最后得到一张由不同条带组成的序列图,然后读出待测 DNA 片段的碱基序列。5′端的第一个碱基 G 读不出来,可以通过测定互补链的序列测出这个碱基。

这一方法的基本步骤概括如下:

1. 先将 DNA 的末端之一进行标记　化学法测定 DNA 序列需要制备各单侧末端标记的待测 DNA 片段,常用的方法是利用 DNA 聚合酶 Ⅰ 大片段,对限制性内切酶产生的 3′凹进末端进行[α-^{32}P]dATP 标记,这种方法能够直接制备单侧末端标记的 DNA,但要求待测 DNA 满足以下条件之一:DNA 两端一侧为 3′凹进末端,另一侧为 3′突出末端或平末端;两侧均为 3′凹进末端,但末端单链结构不同。因此,可以选择不同的[α-^{32}P]dNTP 实现单末端标记。

2. 在多组互相独立的化学反应中分别进行特定碱基的化学修饰。

3. 在修饰碱基位置化学法断开 DNA 链。

4. 聚丙烯酰胺凝胶电泳将 DNA 链按长短分开。

5. 根据放射自显影显示区带,直接读出 DNA 的核苷酸序列。

(三) 图谱的识读

化学法测序的电泳放射自显影图谱的识读比末端终止法复杂一些,每组测序图谱为 4 或 5 条垂直的阶梯带。读片时从胶底部一个个地向顶部读:G+A 和 T+C 两列中含有所有相差一个碱基的 DNA 片段,如果在 G+A 中出现一条带,就看 G 列中是否有同样大小的带,如果有即为 G 碱基,若没有则为 A 碱基;同样,在 T+C 中出现条带,就检查 C 列中有无同样大小条带,有即为 C,无则为 T。如果做了 A>C 反应,在 A>C 列中出现比较深的带时,可以帮助确定其为 A 碱基。

(四) 化学测序法的特点

化学法测序通常都用^{32}P 标记的 DNA,与用^{35}S 标记的末端终止法比较,它显出的条带比较宽且有扩散,这就限制了化学法对较大片段条带的分辨能力,一般在一块电泳胶上最多能读出 200～500 个核苷酸序列。另外,化学法对标记探针及电泳回收 DNA 片段的纯度有比较高的要求。化学测序法已不常用。

三、DNA 序列分析的自动化

Sanger 双脱氧链终止法和 Maxam-Gilbert 化学修饰法都需要使用放射性核素作为标记物,对操作人员有辐射危害,而且放射性材料在保藏、处理和运输方面也相当麻烦。此外,这两种 DNA 序列分析法还存在操作步骤繁琐、效率低、速度慢等缺点,特别是判断测序结果的读片过程,既花时间又乏味。因此,自 DNA 序列分析技术问世不久,就有许多科学家开始致力于 DNA 分析自动化方面的研究。

(一) DNA 自动测序的基本原理

自动荧光序列测定是用染料标记的双脱氧核苷酸作为 DNA 延伸的终止物。DNA 序列

分析自动化包括两个方面：一是“分析反应”的自动化，另一方面则是“读片过程”的自动化。现在大多采用四甲基若丹明(tetramethylrhodamine)作为发光剂，预先标记引物DNA。带有这种引物的DNA片段能在激光诱导下发出荧光，从而取代放射性核素标记。如图11-3所示，由激光发射器产生的激光束，通过精密的光学系统后被导向凝胶表面的检测区垂直射向凝胶，遇到经过检测孔的DNA片段后激发荧光发色基团发射出具特定波长的荧光。这些荧光通过聚焦透镜集中后传给滤光镜/棱镜组件，区分四种碱基产生的不同波长，经成像透镜后由高灵敏度的(CCD)相机分段收集信号，传送到计算机上进行分析处理。

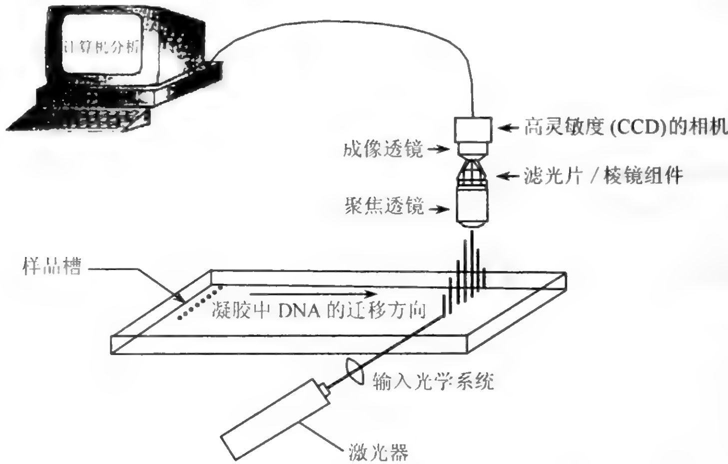


图11-3 高速自动DNA测序仪的结构及工作原理

(二) DNA自动测序的步骤

按照标准的双脱氧链终止法进行测序反应，用聚丙烯酰胺凝胶做电泳分离，在电泳凝胶的侧面，固定上一个激光通道小孔，在凝胶板的上面装上一套荧光信号接收器。电泳过程中，当DNA条带在电场的作用下，经过激光通道小孔时，带有荧光剂标记的DNA在激光的激发下便产生了荧光。荧光感受器收到这种荧光信号后，通过信号转换器转换为电信号，再输入到数据处理系统，最后把所测得的序列直接打印出来(图11-4)。

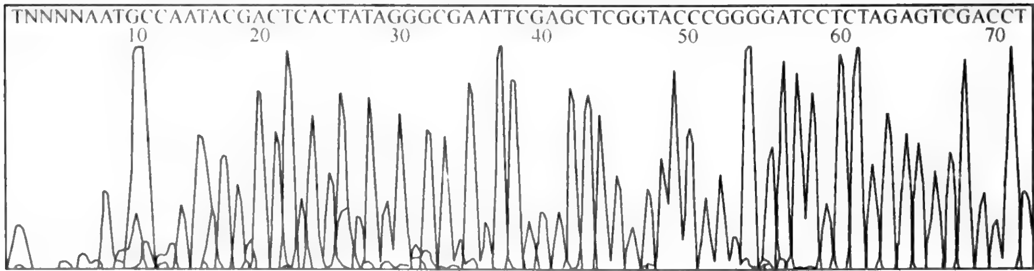


图11-4 自动测序仪利用荧光染料获得的图谱(N代表未测出的碱基)

（三）DNA 自动测序的优点

用染料标记的终止物代替正常 ddNTP 和放射性标记的 dNTP 有三个优点：第一，避免使用放射性物质，加速测序反应；第二，所有四种终止反应可使用四种不同的染料在同一管中进行，一种颜色代表一种 ddNTP，减少了试管和吸头的用量；第三，用不同的染料，所有的反应可在一个泳道上进行凝胶电泳，条带通过它们对扫描仪上不同波长光的吸收能力不同而识别。新式测序仪结合了 *Taq* DNA 聚合酶测序法和染料标记 ddNTP 的优点，是新式自动序列测定的基础。

四、其他的 DNA 测序方法

随着分子生物学的飞速发展，尤其是近年来人类基因组计划（HGP）及后基因组研究的提出，人们需要更加高效而快速的核酸序列分析手段。自从 20 世纪 70 年代末期 DNA 测序技术问世以来，人们已经做了大量重要的改进，但从根本原理上创新的则只有杂交测序法（sequencing by hybridization，SBH）一种。

（一）杂交测序

近年来，一种全新的 DNA 测序方法——杂交测序（sequencing by hybridization，SBH）及基于 SBH 发展起来的 DNA 芯片（DNA chips）或叫基因芯片（gene chips）的技术应运而生。SBH 最初是由英国、前南斯拉夫和俄罗斯的四个科学家小组于 20 世纪 90 年代初同时提出来的，它是利用一组已知序列的寡核苷酸短序列做探针，同某一特定的较长靶 DNA 分子进行杂交，通过计算机分析从而测定其核苷酸序列。DNA 芯片是一种通过杂交测定未知 DNA 序列的新技术。

1. 杂交测序的原理　在一个玻璃或硅片上合成大量已知序列的寡聚核苷酸片段（作为探针），这些探针一端固定在固体基质上，另外一端是游离的。它们在硅片上有规律地排列着，每个特定位置上探针的序列都是已知的。假如有一个 DNA 片段需要测序，我们可以把它的单链形式用荧光进行标记，然后与硅片上的 6 万多种探针进行分子杂交，在荧光显微镜下观察杂交结果。如果某个探针与待测 DNA 某个部分的序列是完全互补的，待测 DNA 分子就被结合到硅片上，这个探针所在的位置就会发出荧光。如果一段较短的 DNA 探针能够同较长的靶 DNA 片段杂交，并形成完全的双链结构，我们便据此推断在靶 DNA 上存在着相应的互补序列。如果将一种核苷酸顺序为 5′-AGCCTAGCTGAA-3′ 的 12-mer 的靶 DNA，与一组完全随机的 8-mer 寡核苷酸探针混合杂交，在总数为 $4^8 = 65\ 536$ 种的 8-mer 探针群体中，仅有 5 种会与靶 DNA 形成完全互补的双链分子。根据这 5 种发生了完全杂交作用的 8-mer 寡核苷酸探针之间重叠序列的线性关系，便可推算出这段 12-mer 的靶 DNA 分子的核苷酸顺序（图 11-5）。

2. DNA 杂交测序的步骤　DNA 杂交测序实质上包括两个主要的步骤，首先是将测定的靶 DNA 分子同一组已知其核苷酸顺序的寡核苷酸探针进行杂交，然后与那些能够同靶 DNA 形成完全的双链杂合分子的寡核苷酸探针做比较分析，并据此推算出靶 DNA 的核苷

酸序列。

(1) DNA 测序芯片的制备:根据待测 DNA 片段的长度选择寡核苷酸探针长度,将该长度所有排列组合的寡核苷酸固定在芯片载体上。一种方法是在芯片载体上直接原位合成,另一种方法是将事先合成好的探针通过喷射或直接接触固定在带有特定活化基团的芯片载体上。

(2) 将待测定的靶 DNA 分子进行荧光标记。

(3) 将带有荧光标记的靶 DNA 分子与 DNA 测序芯片上的探针进行杂交,杂交后进行清洗。

(4) 使用激光共聚焦扫描仪对结果进行荧光检测,得到杂交结果图。根据能够同靶 DNA 完全杂交的寡核苷酸探针之间的碱基重叠关系,通过专门的计算机分析软件处理得出待测 DNA 分子的核苷酸序列。

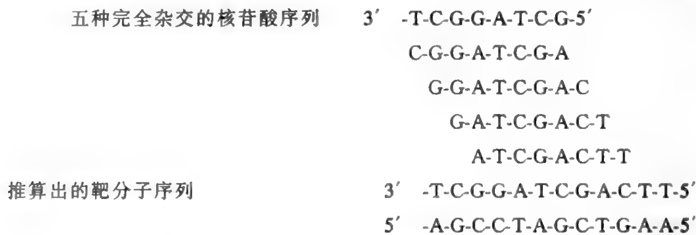


图 11-5　最简单的杂交测序原理

3. 杂交测序的特点

(1) DNA 杂交测序法不仅避免了传统测序法必不可少的操作繁琐的凝胶电泳,而且既不需要使用价格昂贵的核酸酶,也不需要进行复杂的化学反应,被公认为是一种有发展潜力的适合于大规模 DNA 测序工作的新方法,特别适用于 DNA 自动测序。

(2) 由于杂交反应的平行性使得它能在一次实验中自动、同时、快速、敏感地检测数千条序列,而且获得的序列信息高度特异、稳定,它被誉为遗传信息分析方面革命性的里程碑,该技术在病原检测方面也将得到广泛应用。

(3) SBH 技术的效率随着微阵列中寡核苷酸数量与长度的增加而提高,但微阵列中寡核苷酸数量与长度的增加则提高了微阵的复杂性,降低了杂交准确性。邻堆杂交技术(contiguous stacking hybridization,CSH)弥补了 SBH 技术存在的弊端,CSH 技术的应用增加了微阵列中寡核苷酸的有效长度,加强了序列准确性,可进行较长的 DNA 测序。

(二) PCR 测序

PCR 测序有以下优点:①模板需要量小;②方法简便,操作易于标准化、自动化;③测序效率高,准确,在短时间内即可完成。

1. PCR 直接测序(direct sequencing, DS)　PCR 直接测序是指对 PCR 产物进行的直接序列分析,而不需要先将 DNA 待测片段克隆于测序载体上,这不仅大大地简化了操作步骤、节省大量的人力和物力,而且可实现自动化操作,加之新的荧光检测技术的应用,使测序的效率大大提高。

采用 PCR 循环直接测序时,应首先将扩增产物转化为单链测序模板,目前常用的转化方法为不对称 PCR(asymmetric,PCR);此外,还有磁珠俘获法、外切酶消化法和基因组扩增转录同步测序法(genomic amplification with transcript sequeucing,GAWTS)。PCR 直接测序的最新进展是将双脱氧链终止法与 PCR 技术相结合进行测序,在 PCR 反应体系中同时将 ddNTP 加入,并利用放射性核素或荧光素标记的引物引导扩增后,将模板的扩增与测序同时进行,其使用的模板量小且不需分离单链。

2. PCR-寡核苷酸探针斑点杂交(PCR-ASb) 如果一个基因的突变部位、性质经测序分析已经阐明,即可用该方法直接检测突变。该方法的原理即利用人工合成的寡核苷酸片段作为探针,与经 PCR 扩增获得的靶 DNA 进行杂交,在严格控制杂交条件的前提下,通过斑点杂交或其他类型的杂交来检测 PCR 产物中有无相应突变,即探针与靶 DNA 片段之间只要有一个碱基不配对或错配,就能检出。

(1)探针:探针一般合成两个,一个为正常序列,另一含有突变位点,前者与正常靶 DNA 完全互补,后者只与突变 DNA 互补,一般长度为 19 bp,该探针称之为等位特异寡核苷酸探针(allele specific oligo nucleotide probe,ASO 探针)。

(2) 探针的标记:最先应用的标记物为放射性物质如^{32}P、^{35}S,但由于其来源、半衰期及放射性危害而不能普遍应用。利用 PCR 技术可获得足够量的靶 DNA 片段,因此非放射性核素标记的探针亦可获得非常满意的结果,使 ASO 探针使用起来更加简便,而且无放射性核素半衰期的限制,操作亦更安全,适合于临床应用。目前已用于 PKU、β 地中海贫血的基因突变检测。

(3) 杂交类型:PCR 技术的出现,从根本上解决了靶 DNA 的来源问题,使人们能在很短的时间内获得足够量的靶 DNA,传统的 PCR-ASO 技术主要是将 PCR 产物固定于杂交膜上,然后用不同的标记探针检测,也可将已标记好的探针预先固定于杂交膜上,再使之与 PCR 产物结合,检测有无完全互补的 DNA 产物。

3. 测序引物设计与要求 由于在测序 PCR 扩增反应中,所设计的引物对测序 PCR 的结果影响比较大,因此,提出一些设计引物的原则,以供参考。

(1) 引物长度:≥18/20 bp。

(2) 引物应避免 3 个或 3 个以上相同碱基重复序列的出现,特别是要避免 3 个或 3 个以上的连续 Gs。

(3) 在测序 PCR 反应中,T_m 高于 55 ℃ 的引物其结果往往优于具有较低 T_m 值的引物。

(4) GC 含量:40%~60%。GC 含量低的引物,则有必要将引物序列延长到 18 个碱基以上,以确保 T_m 值高于所推荐的低限值 55 ℃。

(5) 引物自身不具有回文结构,以避免引物自身形成发夹结构。如果引物中存在回文结构,则有可能引起测序 PCR 的无效引发,从而影响测序 PCR 结果。

(6) 引物设计应避免引物自身二聚体或引物间二聚体的形成。

(7) 尽可能通过计算机证实对于载体和插入 DNA 片段与所设计的引物的结合位点是惟一的,特别是在 3′末端的 8~10 bp。

(8) 不能设计简并性引物。在测序 PCR 中,简并性引物只能导致测序 PCR 失败。

(9) 由于对多聚核苷酸和短串联重复序列的测序比较困难,因此,设计引物时还要尽量避免与此区域接近或可与之发生退火反应的引物。

(三) 单分子荧光测序

单分子荧光测序是一种快速 DNA 测序法,是利用单分子操作技术直接读取 DNA 碱基序列的方法,与传统的荧光测序法相比,这种方法可大大提高速度。目前,单分子荧光测序技术还没有完全成熟,主要问题是用于识别单个碱基的单分子光谱技术还没有过关,利用隧道扫描显微镜和原子力显微镜直接读取 DNA 分子碱基序列的研究正在进行之中,近期内有可能取得突破。

单分子荧光测序的主要过程如下:第一步,取一条大约有 5 万个碱基的单链 DNA 分子,把它的一端用化学方法连接在一个非常微小的塑料球上,DNA 分子就会缠绕在塑料球上。第二步,在一张类似激光唱片的圆盘上铺一层很薄的液体薄膜,然后用激光光钳把塑料球放在这张光盘的液体膜上进行移动,DNA 分子就会在后面被拖着展开,就好像一只船拖着一条绳子快速前进,把绳子拉展一样。第三步,让拉展的 DNA 分子与一种核酸外切酶结合,这种核酸外切酶可以和 DNA 的游离末端结合,然后逐个把 DNA 的碱基切割下来。第四步,用单分子光谱技术逐个识别并且读出碱基即达到了测序的目的。

(四) 流式细胞仪测序

其原理是以 DNA 单链为模板,合成互补链时的核苷酸分别用不同的荧光染料标记,然后连同核酸外切酶置于石英毛细管内。被酶切而分解下来的核苷酸一个一个地通过流式细胞仪的检测系统,借助于不同的荧光而直接读出核苷酸的序列。

(五) 扫描隧道显微术(scanning tunneling microscopy,STM)

这项技术的基本原理是利用量子隧道效应,将原子线度的极细针尖在接近样品的表面处扫描,通过监测样品与针尖间隧道电源随距离的变化而得到样品表面的形象,对 DNA 样品则可区分出 DNA 糖-磷酸骨架上 4 种碱基的差别。现在最新的一种称为分子激发显微术,扫描针尖与样品之间的能量交换与调制的灵敏度,在分析 DNA 时,如标记上可以区别的金属,则可进一步提高不同碱基的分辨率。

(六) X 射线成像术

积极发展强烈的激光 X 射线源和高质量的 X 射线光学仪器,使 X 射线的成像能力足以在空间上分辨不同的碱基。在发展该项技术时,要考虑消除与防止射线对 DNA 的损伤效应。另一项技术称为准晶体成像术,它是在用 PCR 技术扩增一段 DNA 分子时,以重金属离子标记一种碱基,只要 1000 万个标记的、完全伸展的 DNA 分子就可作为分析样品。当用激光 X 射线作分析时,金属标记物在 DNA 片段中的位置可从散射数据中加以确定。每一 DNA 片段只需测定 A、T、G、C 四种标记物的位置后加以汇总,便可获得核苷酸的排列次序。

第三节 DNA 测序中的常见问题

无论手工测序还是测序仪自动测序,都存在下列问题:

(一) 测序结果信号弱,无法正确读取序列

常见原因是模板量不够,测序模板量主要取决于模板长度和种类。若为 PCR 产物测序,一般要求 100 ng/kb,而质粒 DNA 可稍多,为 100～200 ng/kb。

(二) 测序信号模糊,信噪比显著升高

可能是模板 DNA 污染,比如 PEG、EDTA、RNA 或其他 DNA 导致酶活性受抑制。

(三) 测序信号提前共同终止,测序反应无法继续进行

当模板 DNA 中有二级结构形成、dNTP 浓度不当或模板浓度过大时,可导致测序反应提前终止。

(四)无正常测序信号

常见原因是引物设计不当或模板 G-C 含量过高。

第四节 测序的意义

人类基因组计划(HGP)的最终目的是要了解人类染色体 10 万多种基因的调控机制,从而进一步阐明人类的整个生命活动。因此,对基因密码的测序分析是一项十分重要的内容。生物体内不同的基因密码顺序是区分不同生物种类的依据,疾病的发生亦将导致生物体内基因密码顺序发生改变。应用各种突变检测技术检测到的基因突变,最后都需用序列分析才能确定突变类型及突变位置,其效率可以达到 100%。因此,DNA 序列分析是检测基因突变最直接、最可信的方法,它不仅可确定突变的部位,而且还可确定突变的性质,现已广泛用于基因突变检测、遗传性疾病诊断、单核苷酸多态性(SNP)研究、基因组重叠序列群(conting)之间的间隙填补(补洞)等方面。

(高立芬)

第十二章　生物芯片技术

Chapter12. Technique of Biochip

生物芯片技术是20世纪90年代中期以来影响最深远的重大科技进展之一，是融微电子学、生物学、物理学、化学、计算机科学为一体的高度交叉的新技术。通过设计不同的探针阵列、使用特定的分析方法可使该技术具有多种不同的应用价值，如基因表达谱测定、突变检测、多态性分析、基因组文库作图及杂交测序(sequencing by hybridization, SBH)等，为后基因组计划时期基因功能的研究及现代医学科学及医学诊断学的发展提供了强有力的工具，将会在新基因的发现、基因诊断、药物筛选、给药个性化等方面取得重大突破，为整个人类社会带来深刻广泛的变革。该技术被评为1998年度世界十大科技进展之一。

一、生物芯片的概念

生物芯片(biochip)是指采用光导原位合成或微量点样等方法，将大量生物大分子比如核酸片段、多肽分子甚至组织切片、细胞等生物样品有序地固化于支持物(如玻片、硅片、聚丙烯酰胺凝胶、尼龙膜等载体)的表面，组成密集二维分子排列，然后与已标记的待测生物样品中的靶分子杂交，通过特定的仪器如激光共聚焦扫描或电荷偶联摄影像机(CCD)对杂交信号的强度进行快速、并行、高效地检测分析，从而判断样品中靶分子的数量。由于常用玻片/硅片作为固相支持物，且在制备过程模拟计算机芯片的制备技术，所以称之为生物芯片技术。

二、生物芯片的分类

生物芯片根据点在芯片上的探针的不同分为基因芯片、蛋白质芯片、糖芯片、细胞及组织芯片、芯片实验室。

1. 基因芯片(Gene chip)　又称DNA芯片(DNA chip)，它是在基因探针的基础上研制出的，它将大量DNA探针分子固定于支持物上，然后与标记的样品进行杂交，通过检测杂交信号的强度及分布来进行分析。

2. 蛋白质芯片　与基因芯片的基本原理相同，但它利用的不是碱基配对而是抗体与抗原结合的特异性即免疫反应来检测。蛋白质芯片构建的简化模型为：选择一种固相载体，能够牢固地结合蛋白质分子(抗原或抗体)，这样形成蛋白质的微阵列，即蛋白质芯片。

3. 糖芯片　DN Wang为研究糖基介导的分子识别及抗感染反应在生物医学研究领域

中的作用，研制了以微生物多糖为靶点的糖芯片，将微生物多糖以非化学结合方式固定到表面修饰的玻璃片上，一张玻片上可固定大量的微生物抗原(20 000 个点)。不同糖结构结合的糖结合物均可用于微阵列的制造。经空气干燥的微阵列可以稳定的长期保存。这种糖芯片可同时检测大量的抗体且只需几微升的血清样本。

4. 细胞和组织芯片　是通过微细加工技术将细胞或组织制备在同一张芯片上，用于药物的筛选或组织基因表达或其他分析用。

5. 芯片实验室　为高度集成化的集样品制备、基因扩增、核酸标记及检测为一体的便携式生物分析系统。它最终的目的是实现生化分析全过程全部集成在一片芯片上完成，从而使现有的许多繁琐、费时、不连续、不精确和难以重复的生物分析过程自动化、连续化和微缩化，属未来生物芯片的发展方向。

另外，根据原理不同，生物芯片可分为元件型微阵列芯片、通道型微阵列芯片、生物传感芯片等新型生物芯片。

三、生物芯片的基本工作原理

生物芯片的工作原理包括芯片的制备、样品的制备与标记、杂交、图像的采集和分析。

(一) 芯片的制备

制备芯片方法也不尽相同，以 DNA 芯片为例，基本上可分为两大类：一类是原位合成(即在支持物表面原位合成寡核苷酸探针)，适用于寡核苷酸；一类是预合成后直接点样，多用于大片段 DNA，有时也用于寡核苷酸，甚至 mRNA。

1. 原位合成　有两种途径，一是原位光刻合成(Affymerix 公司专利技术)，该方法的主要优点是可以用很少的步骤合成极其大量的探针阵列。某一含 N 个核苷酸的寡聚核苷酸，通过 $4\times N$ 个化学步骤能合成出 4^N 个可能结构。例如想要合成 8 核苷酸探针，通过 32 个化学步骤，8 个小时可合成 65 536 个探针。而如果用传统方法合成后点样，那么工作量之大将是不可思议的。同时，用该方法合成的探针阵列密度可高达到 $10^6/cm^2$。

另一种原位合成是压电打印法(piezoelectric printing)，原理与普通的彩色喷墨打印机相似，所用技术也是常规的固相合成方法。不过芯片喷印头和墨盒有多个，墨盒中装的是四种碱基合成试剂。喷印头可在整个芯片上移动。支持物经过包被后，根据芯片上不同位点探针的序列需要将特定的碱基喷印在芯片上特定位置。冲洗、去保护、偶联等则与一般的固相合成技术相同。该技术采用的化学原理与传统的 DNA 固相合成一致，因此不需要特殊制备的化学试剂。每步产率可达到 99% 以上，可以合成出长度为 40～50 个碱基的探针。目前除了 Affymetrix 等公司使用该技术合成探针外，其他中小型公司大多使用合成点样法。

2. 点样法　是将预先通过液相化学合成好的探针，或 PCR 技术扩增 cDNA 或基因组 DNA 经纯化、定量分析后，通过由阵列复制器(arraying and replicating device ARD)或阵列点样机(arrayer)及电脑控制的机器人，准确、快速地将不同探针样品定量点样于带正电荷的尼龙膜或硅片等相应位置上(支持物应事先进行特定处理，例如包被以带正电荷的多聚赖氨酸或氨基硅烷)，再经紫外线交联固定后即得到 DNA 微阵列或芯片。

点样的方式分两种，其一为接触式点样（即打印法），即点样针直接与固相支持物表面接触，将 DNA 样品留在固相支持物上；打印法的优点是探针密度高，通常 $1cm^2$ 可打印2 500个探针；缺点是定量准确性及重现性不好，打印针易堵塞且使用寿命有限；其二为非接触式点样，即喷点，它是以压电原理将 DNA 样品通过毛细管直接喷至固相支持物表面。喷印法的优点是定量准确、重现性好、使用寿命长；缺点是喷印的斑点大，因此探针密度低，通常 $1cm^2$ 只有 400 点。点样机器人有一套计算机控制三维移动装置、多个打印/喷印头、一个减震底座，上面可放内盛探针的多孔板和多个芯片。此外，还可以有温度和湿度控制装置、针洗涤装置。打印/喷印针将探针从多孔板取出直接打印或喷印于芯片上。检验点样仪优劣的指标包括点样精度、点样速度、一次点样的芯片容量、样点的均一性、样品是否有交叉污染及设备操作的灵活性、简便性等。

（二）样品的制备

1. 核酸样品的制备　根据样品来源、基因含量、检测方法和分析目的不同，采用的基因分离、扩增及标记方法各异。为了获得基因的杂交信号必须对目的基因进行标记。标记方法有荧光标记法、生物素标记法、放射性核素标记法等。目前采用的最普遍的荧光标记方法与传统方法如体外转录、PCR、反转录等原理上并无多大差异，只是采用的荧光素种类更多，这可以满足不同来源样品的平行分析。

RNA 样品通常需要首先反转录成 cDNA 并进行标记后才可进行检测(图 12-1，A)。目前标记的方法有多种，均可实现对基因表达的检测，其中间接法(图 12-1，B、C)大大提高了标记效率，而 Affymetrix 公司等采用的体外转录方法可使检测的灵敏度明显提高，增加了低丰度表达基因的检出率(图 12-1，D)。由于检测灵敏度所限，尚难以普通探针对极少量的核酸分子进行杂交和检测，所以需要对样品或后续测试信号进行适当的放大。多数方法需要在标记和分析前，对样品进行适当程度的扩增，例如通过 PCR 方法，以使样品核酸的拷贝数有所提高，达到检测的灵敏度。

除了检测前对样品分子的放大外，通常仍需要有高灵敏度的检测设备来采集、处理和解析生物信息。但亦有不经过对样品的扩增和放大而直接应用特殊处理的探针，例如分支探针技术，可达到较高的检测灵敏度水平。这种方法的原理是，设计具有庞大分支结构的分支核苷酸探针，分支末端以酶标记。这样，经过分支核苷酸与酶的双重放大作用而将标本杂交时极弱的信号转换为较强的化学信号。它的最大优点在于其操作简便，具有较高的灵敏度，同时也可以保证检测结果的特异性。当然，由于不同检测方式的灵敏度不同，对于样品的处理和扩增情况的要求亦有所不同，具体的处理和放大方法仍需根据实际情况进行选择。

2. 蛋白质及其他生物样品的制备　蛋白质芯片在进行检测和分析时，可以将待分析的蛋白质样品用荧光素或其他物质进行标记，然后与生物芯片上的生物大分子进行相互作用，最后依据标记物质的不同，采取相应的检测方式采集和分析样品和芯片上生物大分子相互作用的结果。对于非核酸类的生物大分子，存在的问题是有时不便于对其进行扩增和放大，因为其他生物大分子的结构相对比较复杂不能进行简便的克隆或扩增。所以，这就向检测的灵敏度提出了更高的要求。其他生物大分子的检测和分析类似于蛋白质分子。例如核酸与蛋白质的相互作用、配体间的相互作用等。

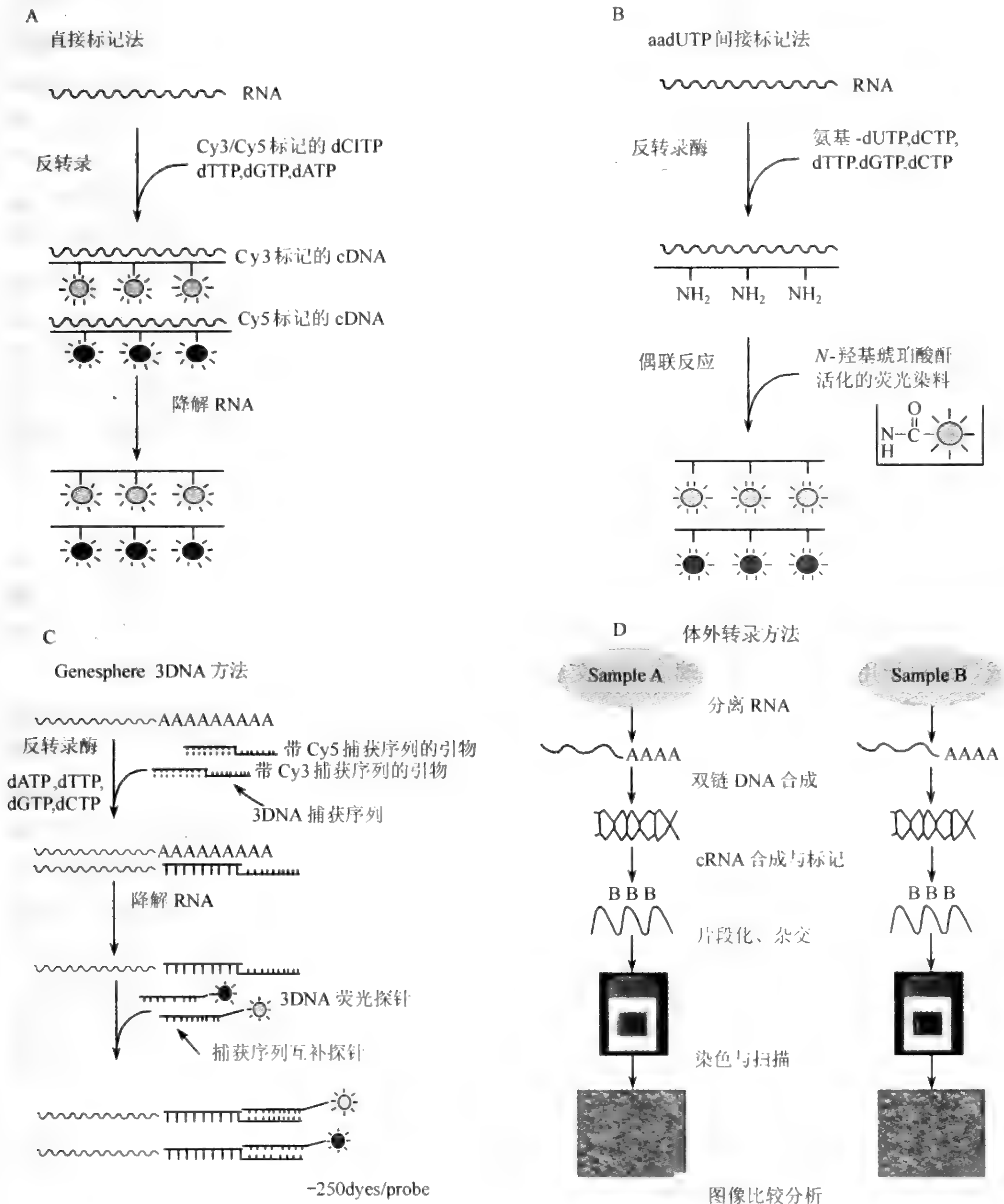


图 12-1　样品的制备

A. 直接标记方法；B. 间接标记方法，此方法目前经常采用；C. 3DNA 间接标记方法（Yu Jetal, Mol Vis. 2002 Apr 26; 8: 130～137）；D. AFFYMETRIX 公司的体外转录标记方法

（三）杂交或反应

互补杂交要根据探针的类型、长度以及研究目的来选择优化杂交条件　如用于基因表

达检测，杂交时需要高盐浓度、高样品浓度、低温和长时间（往往要求过夜），但严谨性要求则比较低，这有利于增加检测的特异性和低拷贝基因检测的灵敏度；若用于突变检测，要鉴别出单碱基错配，需要在短时间内（几小时）、低盐、高温条件下高严谨性杂交。多态性分析或者基因测序时，每个核苷酸或突变位都必须检测出来，通常设计出一套四种寡聚核酸，在靶序列上跨越每个位点，只在中央位点碱基有所不同，根据每套探针在某一特定位点的杂交严谨程度，即可测定出该碱基的种类。杂交反应还必须考虑杂交反应体系中盐浓度、探针 GC 含量和所带电荷、探针与芯片之间连接臂的长度及种类、检测基因的二级结构的影响。有资料显示，探针和芯片之间适当长度的连接臂可使杂交效率提高 150 倍。连接臂上任何正或负电荷都将减少杂交效率。由于探针和检测基因均带负电荷，因此影响它们之间的杂交结合，为此德国癌症研究院的 Jorg Hoheisel 等提出用不带电荷的肽核酸（PNA）做探针效果更好。虽然 PNA 的制备比较复杂，但与 DNA 探针比较有许多优点，如不需要盐离子，因此可防止检测基因二级结构的形成及自身复性。由于 PNA-DNA 结合更加稳定和特异，因此更有利于单碱基错配基因的检测。

（四）图像的采集和分析

当生物芯片和样品探针杂交完毕后，就需要对杂交结果进行图像采集和分析。一般膜芯片的杂交都用放射性核素 ^{32}P、^{33}P 作标记，其信号的检测需通过传统的磷光成像系统来完成。而对于用荧光标记的玻璃芯片杂交后的检测，则需要用专门的荧光芯片扫描仪。

1. 磷感屏成像系统（cyclone storage phosphor system）　Cyclone 磷屏成像系统为美国 Packard 公司生产的第一台集高分辨率、高灵敏度和 5 个数量级的线性范围于一身的计算机控制数字化自动放射成像分析系统，由于其使用方便、快捷、自动化程度、分辨率、图像清晰度均很高，既可定位亦可定量，目前已广泛应用于核医药学、细胞与分子生物学、生物化学、药理学、基因工程学、药动学、放射免疫及受体免疫等多方面实验研究。工作原理在于：放射性核素标记的杂交结果在磷屏上曝光，曝光过程 ^{32}P 等核素核衰变同时发射 β 射线，首先激发磷屏上分子，使磷屏吸收能量的分子发生氧化反应，以高能氧化态形式储存在磷屏分子中。激光扫描磷屏，激发态高能氧化态磷屏分子发生还原反应，即从激发态回到基态时多余的能量以光子形式释放，从而在 PMT（photomultiplier tube，光电倍增管）捕获进行光电转换，磷屏分子回到还原态。计算机接受电信号，经处理形成屏幕图像，并进一步分析和定量。一般化学发光物质如荧光染料标记样品成像过程与放射性类似。

2. 荧光芯片扫描仪　由于杂交时产生序列重叠，会有成百上千的杂交点出现在图谱上，形成极为复杂的杂交图谱。序列重叠虽然可为每个碱基的正确读出提供足够的信息，可提高序列分析的可靠性，但同时信息处理量也大大增加了。一般说来，这些图谱的多态性处理与存储都由专门设计的软件来完成，而不是通过对比进行人工读谱。用计算机处理即可给出目的基因的结构或表达信息。扫描一张 $10cm^2$ 的芯片大概需要 2～6min 的时间。目前专用于荧光扫描的扫描仪根据原理不同大致分为两类：一是激光共聚焦显微镜的原理，是基于 PMT 的检测系统；另一种是 CCD（charge-coupled-devices，电荷耦合装置）摄像原理检测光子。CCD 一次可成像很大面积的区域，而以 PMT 为基础的荧光扫描仪则是以单束固定波长的激光来扫描，因此需要激光头，或者将目的芯片的机械运动来使激光扫到整个芯片，

这样就需要耗费较多的时间来扫描。但是CCD有其缺点：目前性能最优越的CCD数字相机的成像面积只有16mm×12mm(像素为10μm)，因此要捕获到整个芯片(面积20mm×60mm)的话，需要数个数码相机同时工作，或者也可以以降低分辨率为代价来获得扫描精度不是很高的图像。由于灵敏度和分辨率较低，比较适合临床诊断用。

ScanArray利用其专利的激光共聚焦光学系统，通过计算机控制，对生物芯片的荧光杂交信号进行全自动的扫描采集，并通过分析软件对数据结果进行定量分析。

扫描后的图像还需要进一步的处理，这要求一定的软件支持。现有的分析软件包括：Biodiscovery的ImaGene系列，Axon Instruments的GenePix系列，GSI的QuantArray等。

(五) 基因芯片上各个靶点荧光信号的分析原理

用激光激发芯片上的样品发射荧光，严格配对的杂交分子其热力学稳定性较高，荧光强；不完全杂交的双键分子热力学稳定性低，荧光信号弱(不到前者的1/35～1/5)，不杂交的则无荧光。美国GSI Lumonics公司开发出的专业基因芯片检测系统(ScanArray系列)，采用激光共聚焦扫描原理进行荧光信号采集，由计算机处理荧光信号，并对每个点的荧光强度数字化后进行分析。利用QuantArray软件包对扫描的荧光信号进行分析，比较每个克隆在不同组织间表达水平的差别。

(六) 微阵列数据分析

简单来说，微阵列数据分析就是对微阵列高密度杂交点阵图像处理，并从中提取杂交点的荧光强度信号进行定量分析，通过有效数据的筛选和相关基因表达谱的聚类，最终整合杂交点的生物学信息，发现基因的表达谱与功能可能存在的联系。微阵列数据分析主要包括图像分析(Biodiscovery ImaGene 4.0、Quantarray 分析软件)、标准化处理(normalization)、Ratio值分析、基因聚类分析(gene clustering)。

1. 图像分析　激光扫描仪Scaner得到的Cy3/Cy5图像文件通过划格(griding)，确定杂交点范围，过滤背景噪音，提取得到基因表达的荧光信号强度值，最后以列表形式输出。

2. 标准化处理(normalization)　由于样本差异、荧光标记效率和检出率的不平衡，需对Cy3和Cy5的原始提取信号进行均衡和修正才能进一步分析实验数据，标准化处理正是基于此种目的。标准化处理的方法有多种：以一组内参照基因(如一组看家基因)校正微阵列所有的基因、阳性基因、阴性基因、单个基因。

3. Ratio分析(ratio analysis)　Cy3/Cy5的比值，又称R/G值。一般0.5～2.0范围内的基因不存在显著表达差异，该范围之外则认为基因的表达出现显著改变。由于实验条件的不同，此阈值范围会根据可信区间有所调整。处理后得到的信息再根据不同要求以各种形式输出，如柱形图、饼形图、点图等。将每个点的所有相关信息如位标(行及列)、基因名称、克隆号、PCR结果、信号强度、Ratio值等自动关联并根据需要筛选数据。每个点的原始图像另存文件，可根据需要任意排序，得到原始图像的拼图，对于结果分析十分有利。

4. 聚类分析(clustering analysis)　实际是一种数据统计分析。通过建立各种不同的数学模型，可以得到各种统计分析结果，确定不同基因在表达上的相关性，从而找到未知基因的功能信息或已知基因的未知功能。基因聚类分析就是根据统计分析原理，对具有相同统

计行为的多个基因进行归类的分析方法,归为一个簇的基因在功能上可能相似或关联。目前以直观图形显示基因聚类分析结果的程序已有人开发出来,可将抽象的数据结果转化成直观的树形图,便于研究人员理解和分析。聚类分析的方法又分为:

(1) 非监督聚类法(unsupervised clustering):又称配对平均连锁聚类分析(pairwise average-linkage cluster analysis)。该方法是分层聚类的一种形式,非常类似系统发生分析。该方法是基于标准相关系数的计算。K-mean 方法是非监督聚类法的一个变化,目前 Stanford University 的 Botstein 实验室和 NHGRI 的 Trent 实验室都采用该分析方法。

(2) 混合聚类法(hybrid clustering approach):是通过将每一数据点傅立叶变换,寻找那些表达呈周期性变化的基因,比如细胞周期涉及的基因。所谓混合聚类就是先进行非监督聚类,再监督聚类。优点是可以整合以前手工聚类法得到的数据;尤其适合确认细胞周期调控的特征性表达谱。

(3) 神经网络方法(neural network approach):运用自组织图(self organizing maps)并结合非监督聚类法进行聚类。优点是分类标准明确;优化的次序好于其他聚类法;用一种次序风格处理大量数据易于被生物学家接受。

5. 基因表达数据库　基因表达数据库是整个基因表达信息分析管理系统的核心。微阵列数据库起着数据储存和查询、各种相关信息的整合的作用。微阵列数据库可以包含用户的管理信息、原始实验结果(图像文件、信号强度值、背景平均值行列号、基因号等)、各种实验参数(plates/unigene/sets/clusters)、探针相关信息、克隆相关信息[基因名称、基因序列、GenBank accession 号、克隆标志符(IMAGE 和内部)、代谢途径标志符、内部克隆标志符]、分析处理结果、芯片设计相关的资源和数据等。

四、生物芯片的应用

(一) 基因测序

基因芯片利用固定探针与样品进行分子杂交产生的杂交图谱而排列出待测样品的序列,这种测定方法快速而具有十分诱人的前景。Chee M 等用含 135 000 个寡核苷酸探针的阵列测定了全长为 16.6kb 的人线粒体基因组序列,准确率达 99%。Hacia 等用含有 48 000 个寡核苷酸的高密度微阵列分析了黑猩猩和人 *BRCA*1 基因序列差异,结果发现在外显子 11 约 3.4kb 长度范围内的核酸序列同源性在 98.2% 到 83.5% 之间,提示了二者在进化上的高度相似性。

(二) 基因表达水平的检测

用基因芯片进行的表达水平检测可自动、快速地检测出成千上万个基因的表达情况。目前已采用基因芯片对几乎所有的肿瘤进行了表达谱分析。如日本学者采用基因组范围的芯片检测了 27 例慢性髓性白血病的基因表达谱,发现了 150 条表达非常明显上调的基因,其中包括编码细胞周期调节因子、转录活化因子、转录因子蛋白激酶等基因。类似的文章已有上千篇,显示了基因芯片在表达谱研究上的极大的优越性。Schena 等采用拟南芥基因

组内共 45 个基因的 cDNA 微阵列(其中 14 个为完全序列,31 个为 EST),检测该植物的根、叶组织内这些基因的表达水平,用不同颜色的荧光素标记反转录产物后分别与该微阵列杂交,经激光共聚焦显微扫描,发现该植物根和叶组织中存在 26 个基因的表达差异,而参与叶绿素合成的 *CAB*1 基因在叶组织较根组织表达高 500 倍。Schena 等用人外周血淋巴细胞的 cDNA 文库构建一个代表 1046 个基因的 cDNA 微阵列,来检测体外培养的 T 细胞对热休克反应后不同基因表达的差异,发现有 5 个基因在处理后存在非常明显的高表达,11 个基因中度表达增加和 6 个基因表达明显抑制。该结果还用荧光素交换标记对照和处理组及 RNA 印迹方法证实。在 HGP 完成之后,用于检测在不同生理、病理条件下的人类所有基因表达变化的基因组芯片为期不远了。

(三) 生物芯片技术在药物研发中的应用

1. 生物芯片在药物靶点发现与药物作用机制研究中的应用　药物靶点发现与药物作用机制研究是生物芯片技术在药物研发中应用最为广泛的一个领域。在药物靶点发现和药物作用机制研究中所使用的生物芯片主要是指 DNA 芯片。使用 DNA 芯片可以对研究者感兴趣的基因或生物体整个基因组的基因表达进行测定。DNA 芯片在药物靶点发现与药物作用机制研究中的应用具体表现在以下几个方面:

(1) 比较正常的不同组织细胞中基因的表达模式:基因的表达模式给它的功能提供了间接的信息。一些药物的靶点是在整个身体中分布广泛的蛋白质,这类药物的不良反应往往比较大。而选择只在特异组织中才表达的蛋白质作为药物筛选的靶点,可以减少药物对整体产生的不良反应,因而更引起人们的关注。例如骨质疏松症(osteoporosis)与破骨细胞(osteoclasts)的功能有关,破骨细胞可以破坏并吸收骨质,当骨质的形成与破坏出现不平衡的时候,就会导致骨质疏松症。如果破骨细胞的功能得到抑制,那么就可以控制骨质疏松症的发生和发展。利用已有的人类 EST 序列和 DNA 芯片技术,可以容易地得到只在破骨细胞中进行表达的基因如组织蛋白酶 *k* 的基因,它编码半胱氨酸蛋白酶。以组织蛋白酶 *k* 基因作为靶标,筛选对它有抑制作用的药物,就有可能得到治疗骨质疏松症的药物。

(2) 研究正常组织与病理组织基因表达差异:正常组织在病变的过程中,往往伴随着基因表达模式的变化。基因表达水平的升高或降低,可能是病变的原因,也可能是病变的结果。若基因表达的变化是病变的原因,则以此基因为靶点的药物就可能逆转病变;若基因表达的变化是病变的结果,则以此基因为靶点的药物就可能减轻病变的症状。DNA 芯片技术可以在病理组织与正常组织之间一次比较成千上万个基因的表达变化,找出病理组织中表达异常的基因。Heller 等提取正常及诱发病变的巨噬细胞、软骨细胞系、原代软骨细胞和滑膜细胞的 mRNA,用包含细胞因子、趋化因子、DNA 结合蛋白及基质降解金属蛋白酶等几大类基因的 cDNA 芯片进行筛选,发现了数种变化明显的基因。其中除了有已知与类风湿关节炎有关的 TNF、IL-1、IL-6、IL-8、G-CSF、RANTES、VCAM 的基因外,还有编码基质金属弹性蛋白酶 HME、IL-3、ICE、趋化因子 Groα 等的基因。而以前认为金属弹性蛋白酶只存在于肺泡巨噬细胞和胎盘细胞中。弹性蛋白酶可以破坏胶原纤维及组织基底膜层,它在类风湿关节炎病理组织中的出现,为治疗该病提供了新的药物靶点。

利用 DNA 芯片来寻找疾病相关基因的策略尤其适用于病因复杂的情况。例如,恶性

肿瘤的发生常常是多基因共同作用的结果,DNA 芯片技术在肿瘤细胞基因表达模式及肿瘤相关基因发掘中具有重要的作用。Wang 等将一些看家基因、细胞因子及受体基因、细胞分裂相关基因及其他一些癌基因共 5766 个基因的 cDNA 探针固定在芯片上,对正常卵巢组织及卵巢癌组织的 mRNA 进行分析,发现二者之间 30% 基因表达相差 2 倍以上,9% 相差 3 倍以上,其中上调较为明显的有 CD9。上皮糖蛋白(epithelial glycoprotein)、p27 及 HE 蛋白激酶抑制物等。这些结果不仅进一步证实了以前用其他方法获得的结果,还提供了一些新的信息。再如,Kapp 等用包含 950 个 DNA 探针的 DNA 芯片分析比较霍奇金病细胞系 L428 及 KMH2 与 EB 病毒永生化的 B 淋巴细胞系 LGL-GK 的基因表达谱,发现霍奇金病源的细胞系中白细胞介素-13(IL-13)及白细胞介素-5(IL-5)表达异常增高;用 IL-13 抗体处理霍奇金病源的细胞系可显著抑制其增殖。此发现提示,IL-13 可能以自分泌形式促进霍奇金病相关细胞增殖。IL-13 及其信号转导途径可能成为霍奇金病治疗及药物筛选的新靶点。

(3) 建立模式生物细胞中的基因表达模型:采用模式生物细胞进行试验,条件容易控制,对模式生物基因表达的研究将启发人们发现和确认新的药物作用靶点。目前,已有多种模式生物(如酵母)的基因组计划已经完成。酿酒酵母(*Saccharomyces cerevisiae*)就是一种可用来进行药物筛选的较为理想的模式生物。它是真核生物,而且基因组已全部测序,细胞繁殖快,易于培养,与哺乳动物细胞有许多共同的生化机制。现在已经发现,在酵母细胞中存在许多与人类疾病相关的基因。如人类 Werner 综合征表现出早熟的特征,其细胞生活周期变短。人类与此疾病相关的基因与酵母中编码 DNA 解旋酶的 *SGS*1 基因极为相似。Botstein 等得到了 *SGS*1 基因突变的酵母菌株,此突变菌株生活周期变短,细胞的表型特征与患有 Werner 综合征人细胞相似。已有一些研究小组根据公布的酵母基因组序列,用 PCR 方法扩增了酵母 6000 多个开放阅读框(open reading frame,ORF)片段,制成 DNA 芯片,在整个基因组的范围内对酵母的基因表达进行检测。

(4) 建立病原体基因的表达模型:由于病原体的基因组规模相对较小,可用包含其全部基因的 DNA 芯片鉴定那些对人产生毒害作用的基因。异烟肼(isoniazd,INH)是治疗肺结核的常用药物,其治疗结核病的机制是它阻断了分枝菌酸的生物合成途径。Wilson 等根据已测序的肺结核杆菌基因组序列,用 PCR 方法扩增了 3834 个 ORF(占全部 ORF 的 97%),固化在玻片上,制成检测肺结核杆菌基因表达的 DNA 芯片。用 INH 处理敏感菌株,发现除了生化途径已清楚的与分枝菌酸合成相关的一些基因转录水平发生变化外,还发现 *EfpA* 基因的表达也被诱导发生了变化,推测 *EfpA* 基因也参与了分枝菌酸的生物合成,而 *EfpA* 基因只在分枝菌属中的致病菌中才存在,故 *EfpA* 基因可以作为治疗结核病的新靶点。另外,分别用 INH 和 ethionamide 处理敏感菌株,获得了相似的基因表达谱,证实了两者具有相同的作用机制。

(5) 研究药物处理细胞后基因表达变化:药物与细胞(特别是敏感细胞)相互作用,将引起细胞外部形态及内部正常代谢过程的一系列变化。其内部生理活动的变化可集中表现在其基因表达的变化上。通过测定分析药物对细胞的基因表达的影响,可推测药物的作用机制,评价药物活性及毒性,进而确证药物靶点或者发现新的药物靶点。通过 DNA 芯片测定药物诱导的细胞基因表达变化来进行药物筛选与研究,对那些用常规方法很难追踪监测的药物或需要很长时间才能得到药物临床实验结果时,显得尤为有用。通过监测阳性药物处

理前后组织细胞基因表达变化情况可以获得许多十分有价值的信息。首先,经药物处理后表达明显改变的基因往往与发病过程及药物作用途径密切相关,很可能是药物作用的靶点或继发事件,可作为进一步药物筛选的靶点或对已有的靶点进行验证;其次,药物处理后基因表达的改变对药物作用机制研究有一定的提示作用。

细胞周期蛋白依赖型激酶(cyclin-dependent kinas es,CDKs)在细胞增殖中有着重要的作用,是一种抗肿瘤药物的靶点。Gray 等从 2,6,9-三取代嘌呤化合物库中筛选 CDKs 的抑制剂。他们将体外有活性的 2 种化合物 flavopiridol 和 52(化合物代号)与酵母作用,然后用共含有 260 000 个长度为 25 个碱基的一组寡聚核苷酸芯片检测酵母基因表达变化。试验结果表明,几乎所有的已知与细胞周期相关的基因表达下调。虽然 flavopiridol 和 52 体外活性相似,但它们引起的酵母基因表达的变化却有显著的差异,表明二者对酵母的作用途径是不一样的。

鬼臼亚乙苷(etoposide)是 *p*53 活化的拓扑异构酶 Ⅱ 的抑制剂,在临床上作为一种抗肿瘤药物。Wang 等经用鬼臼亚乙苷作用于人成骨肉瘤细胞系 U2-OS 后,接种后不同的时间间隔分别提取细胞 mRNA,用寡核苷酸芯片测定 6591 条 mRNA 表达的变化,发现 62 条 mRNA 表达有变化。通过选取其中 12 条基因做进一步研究,发现有 2 条是已知的 *p*53 调控基因(WAF1/p21 和 PCNA),有两条是新的 *p*53 靶基因,其余的与 *p*53 无关。在实验基础上,他们提出了介导鬼臼亚乙苷诱导细胞凋亡的信号传导途径。此项研究工作使人们又多获得一种抗肿瘤药物的靶点。应用 DNA 芯片还可直接筛选特定的基因文库以寻找药物的作用靶点。

除了 DNA 芯片外,组织芯片、蛋白质芯片和细胞芯片也在药物研究中崭露头角。最近,耶鲁大学的研究小组首次报道了真核生物蛋白质组水平的蛋白质微阵列芯片。他们表达和纯化了酵母的 5800 种蛋白质,并将这些蛋白质点样固定在载玻片上,制备了酵母蛋白质组微阵列芯片。他们使用这种芯片筛选能与特定蛋白质和磷脂相互作用的蛋白质,发现了新的能与钙调蛋白和磷脂相互作用的蛋白质。这种蛋白质微阵列芯片可以用于筛选与蛋白质相互作用的药物,还可以用于检测蛋白质翻译后的修饰。他们的研究成果证实了制作和使用蛋白质组微阵列芯片进行功能分析检测的可行性,并向人们预示了蛋白质微阵列芯片在药物开发领域的广阔应用前景。

2. 生物芯片作为超高通量筛选平台的应用　在过去的十几年里,随着科学的进步以及在巨大的经济利益驱使下,药物筛选技术得到了飞速的发展。在 20 世纪 80 年代中期(高通量筛选形成之初),每天只能筛选 30 种化合物,到 90 年代中期,每天可筛选 1500 种化合物,而如今每天可筛选超过 100 000 个化合物。提高药物筛选的通量,实现超高通量筛选有 2 条途径:一是微型化,一是自动化。生物芯片作为一种新型技术平台,正可满足超高通量筛选微型化和自动化的需要。生物芯片技术应用于超高通量筛选有 2 个发展方向:一是微孔板/微阵列技术,一是微流体芯片技术。

(1) 微孔板/微陈列技术:微孔板技术的发展主要表现在板孔数的增加。目前,使用最多的是 96 孔及 384 孔板,也有人使用 1536 孔、3456 孔,甚至 9600 孔板。如 Oldenburg 等报道了用 9600 孔板(0.2μl/孔)分析系统,以金属蛋白酶为靶,对组合及分离纯化的化合物库进行筛选的结果。虽然随着材料科学和加工技术的发展,微孔板技术有了长足的进步,但其

发展面临着一些不易解决的困难，主要有：微量液体极易蒸发，不适于那些不能用二甲亚砜(DMSO)做溶剂的筛选方法以及受限于当今还不够完善的微量液体分配技术。

微阵列技术是将微孔板技术进一步微型化。最近，哈佛大学的研究人员开发了化合物微阵列芯片，主要用于筛选能与特定蛋白质特异性结合的化合物。他们将玻片表面进行化学处理，使其衍生化产生活性基团，然后将溶于有机溶剂中的化合物用机械手点在经处理的玻片表面，化合物与玻片表面的活性基团反应而被固定于玻片表面，这样就将不同的化合物排布成微阵列，固定在玻片表面，制成化合物微阵列芯片。随后将感兴趣的蛋白质进行荧光标记，然后与微阵列芯片上的化合物反应，经清洗后，再进行荧光检测就可以筛到能与这种蛋白质特异性结合的化合物。他们用化合物微阵列芯片进行了原理性实验，其结果表明，使用这种化合物微阵列芯片可以并行、快速地进行大规模的化合物与蛋白质的结合筛选。

(2) 微流体芯片技术：微流体装置的发展已广泛用于生化及细胞的分析。它借用半导体工业中所用的光刻技术将内径在 10～100μm 的通道加工在玻璃或硅片中，利用电动泵和流体的压力来控制皮、纳升级液体的流动。该技术可减少几个数量级的试剂消耗量，并能提高数据质量。它所采用的并行样品处理程序可以获得更高的筛选通量。多种类型的筛选分析方法在微流体芯片上操作的可行性(包括结合分析、酶分析和细胞分析)已得到实验论证。

Capliper Technologies 公司设计的芯片装置，将从微孔板中取样与随后进行的酶抑制剂筛选结合在一起。目前，该芯片装置已制作完成，并通过了测试。测试的化合物通过与芯片相连的硅毛细管注射入芯片中，与酶及荧光酶底物混合以使它们在主反应通道内发生相互作用。酶与荧光底物反应产生基线荧光信号。如果测试化合物抑制酶的活性，信号将会暂时性的降低，通过检测荧光信号的强度就可以测定化合物抑制酶的活性。这种装置每次进样量为 1～10nl，进样周期为 10～30s。利用这个系统，每块芯片可对几千种化合物进行筛选。目前对该系统进行的改进主要集中在提高系统的耐用性和将不同类型的筛选方法移植到这个平台上来。最新设计的微流体芯片将液体分发、化合物稀释与筛选分析三者集成起来，由于减少了被视为瓶颈的芯片使用前的准备步骤，使得该系统更加方便实用。

3. 生物芯片在毒理学研究中的应用　对药物进行毒性评价，是药物筛选过程中十分重要的一个环节。现在毒理学家多采用鼠为模型通过动物实验来确定药物的潜在毒性。这些方法需要使用大剂量的药物，花上几年时间，花费巨大。DNA 芯片技术可将药物毒性与基因表达特征联系起来，通过基因表达分析便可确定药物毒性，使得药物毒性或不期望出现的效应在临床实验前得以确认。用 DNA 芯片可以在一个实验中同时对成千上万个基因的表达情况进行分析，为研究化学或药物分子对生物系统的作用提供全新的线索。该技术可对单个或多个有害物质进行分析，确定化学物质在低剂量条件下的毒性，分析、推断有毒物质对不同生物的毒性可比性。如果不同类型的有毒物质所对应的基因表达有特征性的规律，那么，通过比较对照样品和有毒物质的基因表达谱，便可对各种不同的有毒物质进行分类，在此基础上通过进一步建立合适的生物模型系统，便可通过基因表达变化来反映药物对人体的毒性。

已经有不少研究工作表明，利用 DNA 芯片预测化合物毒性和对毒性物质进行分类是可行的。Waring 等用 15 种已知的肝毒性化合物处理大鼠。这些毒物将对肝细胞造成多种伤害，如 DNA 损伤、肝硬化、肝坏死和诱发肝癌等。从大鼠肝脏中提取 RNA，用 DNA 芯片

做基因表达分析。通过将基因表达结果与组织病理分析和临床化学分析的结果进行比较，发现两者有很强的相关性。该结果表明，DNA 芯片分析是一种可以用来分析药物安全性和对环境毒物进行分类的灵敏度较高的方法。在另一报道中他们用同样的 15 种化合物作用大鼠的肝细胞，再用 DNA 芯片做基因表达分析，结果显示具有相似毒性机制的化合物所获得基因表达谱具有相似性。Gerhold 等给大鼠服用苯巴比妥和地塞米松等药物，使用寡核苷酸芯片，检测了大鼠肝组织中与药物代谢、毒性和能量代谢相关基因的表达。最后，通过分析基因表达变化的结果就可以推测药物代谢与毒性的情况。Bartosiewicz 等则利用 DNA 芯片对环境毒物进行了检测。

4. 生物芯片在药物基因组学中的应用 药物基因组学是在基因组学的基础上研究不同个体对药物反应的差异，以便针对不同的基因型"量身定做"药物，从而将药物的药效充分发挥而不良反应减少到最小。其优点为：①在进入临床试验前，药物基因组学可以通过化合物对基因多态性的影响挑选先导物，从而降低由于药效的不稳定导致的失败几率。②在Ⅰ期临床试验中，个体基因型可以预见基因多态性造成的药动学差异。③由于药物作用靶蛋白的差异反映在基因多态性上，因此在Ⅱ期临床试验中，由个体基因型可以预见基因多态性造成的药效差异，由此来指导Ⅲ期临床试验 ④一旦发现一种可以导致药物作用差异的蛋白质，其他与之相关的蛋白质可作为潜在的药物作用靶。

DNA 芯片可以自动、快速地检测那些可以影响药物效应的基因（如药物代谢酶、药物作用靶和致病因子的编码基因）和决定个人对药物毒性敏感性的基因。Evans 和 Relling 就设想了一种急性淋巴细胞白血病药物基因组芯片，这种芯片上包括了所有可能影响急性淋巴细胞白血病病人对化学治疗反应的基因，借助这种芯片可以根据病人的基因型对病人分群，帮助医生为每个病人选择合适的治疗药物和药物剂量。

5. 生物芯片在药物分析中的应用 生物芯片在药物分析中的应用主要是指采用毛细管电泳芯片/质谱系统对化合物库、血样和尿样中的药物进行分析鉴定。毛细管电泳芯片/质谱系统是指将毛细管电泳芯片和质谱联用的一套装置。毛细管电泳芯片进行样品的分离，而与芯片联用的质谱则有选择地对分离成分进行检测。美国康奈尔大学的 Wachs 等发明了一种微型化的离子喷雾装置。这种装置适合于基于芯片的分离装置、多孔板或带有待测样品残渣的表面联用。这种装置有两种版本，一种称为微型喷雾器，主要与毛细管电泳芯片联用；另一种称为小型喷雾器，它带有伸长的吸样毛细管，可以插入多孔板孔的底部，所以适合与多孔板联用。这种装置可以帮助人们对芯片分离所得样品或多孔板中样品进行质谱检测。来自康奈尔大学同一个实验室的 Deng 等将玻璃制成的毛细管电泳芯片与他们自己设计的微型离子喷雾装置连接，使用 PE 公司的质谱系统做检测，对多种药物的标准和血浆样品进行了分析检测。实验结果证明，使用毛细管电泳芯片/质谱系统对小分子化合物进行快速(30s)的检测是可行的。这套装置可以用来对合成的化合物库和人血浆中重要成分进行分析检测。Ramseier 等用芯片毛细管电泳和凝胶毛细管电泳检测尿中苯丙胺及其类似物，实验结果显示，芯片毛细管电泳的方法具有更快的分高速度和更高的分离效率。Chiem 等则用毛细管电泳芯片检测了血清中的茶碱。

(四) 生物芯片在临床上的应用前景

基因芯片在感染性疾病、遗传性疾病、重症传染病和恶性肿瘤等疾病的临床诊断方面具有独特的优势。与传统检测方法相比,它可以在一张芯片同时对多个病人进行多种疾病的检测,无需机体免疫应答反应期,能及早诊断,待测样品用量小。能特异性检测病原微生物的亚型及变异;可帮助医生及患者从“系统、血管、组织和细胞层次”(通常称之为“第二阶段医学”)转变到“DNA、RNA、蛋白质及其相互作用层次”(第三阶段医学)上了解疾病的发生、发展过程,这些特点使得医务人员在短时间内可以掌握大量的疾病诊断信息,这些信息有助于医生在短时间内找到正确的治疗措施。

众所周知,肿瘤和遗传疾病发生的根本原因是由于遗传物质发生了改变,检测基因突变对于阐明肿瘤及遗传病的分子机制、疾病的早期诊断具有重要意义。目前,Affymetrix 公司已开发出 *p*53 基因芯片,是将已知 *p*53 基因全长序列和已知突变的探针固定在芯片上,这将有助于恶性肿瘤的早期诊断。华盛顿大学的分子生物学系与病理系联合研究了卵巢癌中基因表达谱的变化,他们将 5 766 个基因探针固定于芯片上,其中 5 376 个分别选自卵巢癌、卵巢表面上皮细胞及正常卵巢的 cDNA 文库,另外还有 342 个来自 EST 克隆,包括一些已知确定的管家基因、细胞因子及其受体基因、生长因子及其受体基因、与细胞分裂相关的基因以及新近确定的肿瘤相关基因,找出在卵巢癌组织中过度表达的 30 个由 GenBank 收录的基因,如高表达的有 *CD*9(GenBank 录入号:M38690)、Epithelial glycoprotein (GenBank 录入号:M32306)、*p*27(GenBank 录入号:X67325)等,这证明了利用基因芯片分析复杂生物体系中分子变化的可行性。又如,Hacia 等在 1.28 cm×1.28 cm 的芯片上固定了 9.66×10^4 个长度为 20 nt 的寡核苷酸探针,用于检测乳腺癌基因 *BRCA*1 的 *exon*11(3.45 kb)中所有可能的碱基置换、插入和缺失(1～5 bp)突变。在 15 例患者样品中,发现有 14 例有基因突变,类型包括点突变、插入及缺失等;在 20 例对照样品中均未检出假阳性结果。目前,肝炎病毒检测芯片、地中海贫血、多种恶性肿瘤相关病毒基因芯片等多种芯片已面市,相信不久的将来,会有更多的基因芯片用于临床。

(五) 生物芯片在环境监测、环境保护方面的应用

在环境监测和环境保护上,可以利用生物芯片快速检测污染微生物或有机化合物对环境、人体、动植物的污染和危害。我们都知道,环境的变化会引起细胞发生基因表达谱的变化,DNA 芯片是高效的监测基因表达谱变化的手段。所以,国外许多实验室纷纷探索用 DNA 微阵列技术监测环境污染和提高环境保护技术水平。生物芯片在环境保护方面的应用还可以通过大规模的筛选寻找保护基因,制备防治危害的基因工程药品及能够治理污染源的基因产品。

五、存在问题及发展前景

目前,国外已经有十多家从事 DNA 芯片研究的公司,他们的芯片技术虽然各具特色,应用目的也不尽相同。但要成为实验室研究或临床可以普遍采用的技术仍有一些关键问题

亟待解决，如：①提高基因芯片的特异性；②简化样品制备和标记操作；③增加信号检测的灵敏度；④高度集成化样品制备、基因扩增、核酸标记及检测仪器的研制和开发。上述问题不仅是当前和今后一段时期内国内外基因芯片技术研究的焦点，同时也是基因芯片能否从实验室研究推向临床应用的关键问题。总之，基因芯片技术发展到今天不过短短几年时间，虽然还存在着问题，但其在基因表达谱分析、基因诊断、药物筛选及序列分析等诸多领域已呈现出广阔的应用前景。随着研究的不断深入和技术的更加完善，基因芯片一定会在生命科学研究领域发挥出其非凡的作用。从某种意义上说，我们可以这样认为：基因的结构或种类决定物种；基因的功能或表达则决定生命，即生物的生、老、病、死。基因芯片技术将为我们提供一条认识生命本质的捷径。

（韩金祥　高雪芹）

第三篇　基因工程学的应用

Part Ⅲ. Application of Genetic Engineering

第十三章　基因工程抗体

Chapter13. Genetic Engineering Antibody

自 1984 年第一个基因工程抗体——人鼠嵌合抗体研制成功以来,新型基因工程抗体不断出现,如人源化抗体、小分子抗体和抗体融合蛋白(Fab、单链抗体、双链抗体、双特异抗体)等。另外,噬菌体抗体库与核糖体展示文库技术的出现,使不经免疫即可获得包括人在内的任何一种动物的特异性抗体成为可能。

第一节　基因工程抗体概述

抗体作为疾病预防、诊断和治疗的制剂已有上百年的发展历史。其制备方法经历了多克隆抗体、单克隆抗体,现已发展至第三代的基因工程抗体。

多克隆抗体(多抗,PcAb,polyclonal antibody)是人类有目的的利用抗体的第一步,其制备方法是将天然抗原经免疫动物而获得。多抗的不均一性,限制了对抗体结构和功能的进一步研究和应用。1975 年,Köhler 和 Milstein 首次用杂交瘤技术制备出具有高度均一性的单克隆抗体(单抗,McAb,monoclonal antibody)这是抗体工程发展的第一次质的飞跃,也是现代生物技术发展的一个里程碑。单克隆抗体因其均一性高、纯度大而在疾病诊断、治疗和科学研究中得到了广泛的应用。单克隆抗体具有鼠源性,进入人体会引起机体产生人抗鼠抗体(human anti-mouse antibody,HAMA),导致排斥反应;同时,单抗是完整的抗体分子,其相对分子质量较大,在体内穿透血管的能力差;生产成本高,不适合大规模工业化生产。20 世纪 80 年代,随着分子生物学、分子免疫学的发展,抗体基因结构和功能的研究成果与重组 DNA 技术相结合,产生了基因工程抗体技术。

基因工程抗体(genetic engineering antibody)指利用重组 DNA 技术和蛋白质工程技术,将抗体基因按不同需要进行加工、改造和重新装配,然后经转染导入适当的受体细胞中表达的抗体分子。与单克隆抗体相比,基因工程抗体具有如下优点:

1. 降低免疫原性　通过基因工程技术的改造,可以降低甚至消除人体对抗体的排斥反应。

2. 改善抗体药物动力学　基因工程抗体的相对分子质量较小,可以部分降低抗体的鼠源性,更有利于穿透血管壁,进入病灶的核心部位。

3. 可根据治疗的需要,制备新型抗体。

4. 生产简单,价格低廉　可以采用原核细胞、真核细胞和植物等多种表达方式,大量表达抗体分子,大大降低生产成本。

基因工程抗体的发展经历了一个长期的过程:最初的基因工程抗体主要是利用分子生物学技术,对单克隆抗体进行人源化改造。由于人源化抗体结构复杂、相对分子质量大,无法改善抗体的亲和力,这些缺点限制了其更为广泛的应用。小分子抗体的出现,大大改善了

药代动力学，在诊断和治疗领域得到了广泛应用。组合化学方法在抗体工程领域中应用，诞生了抗体库技术。噬菌体抗体库是目前应用最为普遍的抗体库技术。采用噬菌体抗体库技术筛选抗体避免了动物免疫、细胞融合及克隆筛选的长期、复杂过程，易于制备稀有抗原的抗体，并使制备完全人源化的高亲和力抗体成为可能。在噬菌体抗体库基础上，近几年又发展了核糖体展示抗体库技术。核糖体展示抗体库技术的整个过程均在体外进行，可以构建高容量、高质量的抗体库，代表了抗体工程的未来发展趋势。

第二节　人源化抗体

人源化抗体(humanized antibody)，是为了克服鼠源性单克隆抗体的免疫原性，而通过不同手段，将其进行人源化改造后获得的抗体，其目的是使之和人体自身的抗体分子具有极其相似的轮廓，从而降低免疫原性，逃避人免疫系统识别，避免 HAMA 产生。

抗体人源化改造，主要基于对抗体本身结构的认识。根据氨基酸变化程度，抗体的肽链可以分为可变区(V 区)和不变区(C 区)两部分，其中，可变区包含了与抗原结合的位点，其免疫原性较小，而不变区与抗原结合功能无关，有较强的免疫原性。因而，人源化抗体主要是对可变区或不变区的改造。

抗体的人源化主要有两个基本原则：一是保持或提高抗体的亲和力和特异性；二是降低或基本消除抗体的免疫原性。不管在人源化中采用何种方案进行，都必须遵循这两点，尤其不能丧失抗体特异结合抗原的能力。

一、人-鼠嵌合抗体

由于不变区与抗体的抗原结合功能无关，而且 90％ 的 HAMA 是针对不变区的，因而第一代的人源化抗体设计用人的 C 区替代鼠的 C 区，即保留鼠 McAb 的可变区，并使之与人抗体的恒定区重组，成为嵌合抗体(图 13-1)。

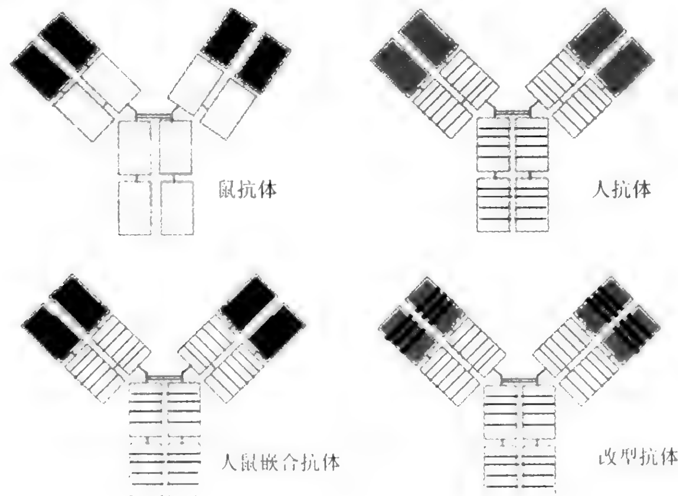


图 13-1　人源化抗体

嵌合抗体的基本原理及过程：

1. V区基因的克隆 尽管Ig V区基因具有复杂的多样性，但Ig重链和轻链V区基因5′及3′端(骨架区及前导肽序列)相对保守，因而可以设计相应的引物通过反转录PCR从分泌某种鼠单抗的杂交瘤细胞中扩增重链及轻链的相应功能区。

2. 表达载体构建 将克隆的鼠单抗功能性V区基因，与人抗体C区基因按一定方式重组，克隆到表达载体中，即可构建鼠-人嵌合的轻重链基因表达载体。然后将嵌合抗体的DNA重组子导入受体细胞。

3. 嵌合抗体的表达 由于嵌合抗体是完整的抗体分子，其恒定区必须经糖基化才能发挥功能。因此，大肠杆菌、噬菌体一般不能作为宿主细胞。酵母菌因为存在糖基化错误的可能性，通常也很少用于表达嵌合抗体。COS细胞是短暂表达嵌合抗体较为理想的细胞，一般在几天内即可装配成功，从而为快速分析嵌合抗体提供了可能。骨髓瘤细胞是表达嵌合抗体最主要的宿主细胞，其中又以SP2/0细胞为主。此外，中国仓鼠卵巢瘤(CHO)细胞和肿瘤浸润性T淋巴细胞等都可以用来表达嵌合抗体。

4. 嵌合抗体表达产物的检测 ELISA法、Western印迹法、RIA以及间接荧光免疫分析法等均可用来检测嵌合抗体。

人鼠嵌合抗体用人的C区取代鼠的C区，最大限度地保留了鼠单抗的特异性和亲和力，较好地解决了鼠源性单抗诱发的HAMA等不良反应，并可延长半寿期，改善药动学。此外，嵌合抗体含有人抗体的Fc段，能有效地与靶细胞上的Fc受体结合，诱导细胞毒性效应。1998年，美国FDA批准抗TNF-α人鼠嵌合抗体用于炎症性肠病，开辟了嵌合抗体用于人体治疗的先河。目前，嵌合抗体已经用于肿瘤、感染、自身免疫病等疾病的治疗，并显示出良好的治疗效果。但是，嵌合抗体中保留了完整鼠抗体的V区结构，使其仍有较强的异源性。而且，嵌合抗体是完整的抗体分子，相对分子质量较大，对肿瘤组织的穿透能力较差，清除也较慢，因而在导向诊断和治疗等方面的应用受到限制。

二、改型抗体

为进一步减少嵌合抗体的鼠源性序列，第二代的人源化抗体在嵌合抗体的基础上进一步将鼠单抗可变区中相对保守的骨架区(framework region，FR)替换成人的FR，仅保留与抗原结合的CDR(complementary determinant region)部位，称为改型抗体(reshaped antibody, RAb)。换言之，RAb是指利用基因工程技术，将人抗体V区中的CDR序列改换成鼠单抗的CDR序列后重构成的一种新的基因工程抗体，又称为重构型抗体。因RAb主要涉及CDR的“移植”，故又称作CDR移植抗体(CDR grafting antibody)(图13-1)。

RAb的构建基于对抗体分子结构的进一步认识。基因序列及X线晶体衍射法的分析显示，抗体分子的V区是由CDR和FR两部分构成。V_H和V_L上的6个CDR折叠成环状，并向外突出形成抗原结合位点，从而决定了抗体的特异性和亲和力。V区中的FR序列基本保守，且在不同的抗体间有较高同源性，同一抗体的FR可适用于不同抗原结合位点的移植。抗体分子的这种结构特性，提供了设计和构建RAb的物质基础。

RAb 的发展经历了两个阶段:

1. 第一代改型抗体　第一代 RAb 是以人骨髓瘤蛋白的已知晶体结构为根据,选择序列已知、并与鼠源性单抗有较高同源性的人 FR 作为框架,并移植人鼠单抗的 CDR,制备除 CDR 序列不同外,其他结构几乎完全相同的治疗用抗体。为保证 RAb 的亲和力,可改变个别氨基酸残基的组成。第一代改型抗体仅通过简单的移植,将鼠单抗的 CDR 与人抗体进行重构,虽然最大限度地减少了异源性序列,但却极有可能降低抗体的亲和力。

2. 第二代改型抗体　为获得具有高亲和力的治疗性抗体,第二代 RAb 的构建应用了人抗体基因库,引进计算机技术模拟抗体分子的立体结构(分子模拟法),并使用了人鼠嵌合 FR。其基本过程为:首先构建人抗体基因库,根据同源性原理选择出与鼠单抗高度同源的人 FR 序列,并通过分子模拟法鉴别出与抗原结合位点关系密切的鼠单抗 FR 上的氨基酸残基,在构建 V 区时把这些关键 FR 残基与 CDR 重组并移植到所选取的人 FR 上,形成鼠人嵌合 FR,其亲水性残基来自人抗体序列,而功能区内的残基及连接 CDR 的残基则来自鼠源抗体序列,从而在保证抗体特异性的同时,获得了高亲和力的 RAb。

RAb 中由于保留了鼠单抗与抗原结合的最小单位(CDR 序列)而保持了原有单抗的特异性,并最大限度地减少了鼠源性。RAb 的产生与发展,使多种特异的鼠源单抗有可能应用于临床治疗,其中包括通过人体免疫应答难以诱生的特异性抗人抗原的抗体,因而具有广泛的应用前景。

第三节　小分子抗体

人源化抗体具有完整的抗体结构,相对分子质量较大,难以穿过血管壁,影响靶部位对其的摄取。特别是肿瘤组织的血供本身就不丰富,对抗体的摄取量就更少。通过基因重组技术,在保持原有抗原结合活性的基础上,把完整的抗体分子改造成小分子抗体,可改善抗体的药动学特性,并适合于临床的应用。小分子抗体根据构建的方法不同,可分为 Fab 抗体、单链抗体、单域抗体和双特异性抗体(图 13-2)。

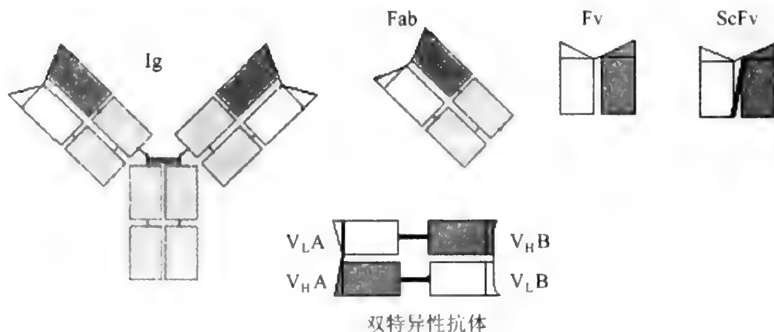


图 13-2　小分子抗体结构示意图

一、Fab 抗 体

抗体的 Fab 段只有完整 IgG 的 1/3，由重链 V 区及 C_H1 功能区与轻链间以二硫键形式连接而成，具有与完整抗体完全相同的抗原结合能力。采用基因工程手段构建 Fab 抗体的原理，是把抗体分子的 V_H 区和 C_H1 功能区的 cDNA 与轻链完整的 cDNA 序列重组并克隆到适当的表达载体中，然后在大肠杆菌等宿主细胞中表达。由于 Fab 抗体不具有 Fc 段，不需要翻译后的糖基化修饰，因而多采用大肠杆菌来表达。此外，Fab 抗体也可在植物细胞及昆虫细胞内表达，但不同的表达类型其表达量有明显的不同。

Fab 抗体的相对分子质量较小，约为 50 kD 左右，体内应用具有较好的渗透性和药动学特征，已被用于多种疾病的诊断和治疗。但多数 Fab 抗体具有较大的鼠源性，人体内反复使用易引起 HAMA 反应。近年，用噬菌体抗体库技术来制备完全人源的 Fab 抗体，有望解决这个问题。

二、单 链 抗 体

单链抗体(single chain of Fv, ScFv)是由 V_H 和 V_L 通过连接肽重组并表达的一种小分子抗体，具有抗原结合能力良好、相对分子质量小、穿透力强、体内循环半衰期短及免疫原性低等特性。在此基础上，可与其他效应分子构建成多种具有新功能的抗体分子，是构建免疫毒素和双特异抗体等的理想而基本的元件。因此，单链抗体在基因工程抗体的研究中占有重要的地位。

单链抗体基因的构建一般采用从杂交瘤细胞提取 mRNA，反转录成 cDNA，通过 PCR 扩增其 V_L 及 V_H 基因，再用人工合成的寡核苷酸序列(即接头)把 V_L 的 C 端与 V_H 的 N 端或 V_H 的 C 端与 V_L 的 N 端连接，即成为单链抗体基因。近年来，将噬菌体表面呈现技术(phage display)应用于单链抗体筛选，可从人外周血淋巴细胞直接提取 mRNA，绕过免疫和细胞杂交融合的过程，从而直接可以构建和筛选单链抗体基因。

设计理想的接头(linker)序列是构建单链抗体的关键。理想的接头需保证 V_H 和 V_L 在表达系统中能等摩尔地产生，不干扰 V_H 和 V_L 的自由折叠，并使抗原结合位点处于适当构型，不引起分子动力学改变，尽可能减少蛋白酶攻击及防止单链抗体的聚集等。目前最常用的接头序列是具刚性结构的 15 肽序列(G1y4Ser)3。此外，还可从已知三级结构的天然蛋白质获得合适的序列，如轻链的 elbow region 序列，藻类蛋白序列等作为接头序列。

单链抗体虽然较好地保持了抗体的特异性，但当抗原含重复抗原决定簇时，其亲和活性可降低 8 倍以上。通过构建双价 ScFv，可改善 ScFv 的亲和活性。将噬菌体表面呈现技术应用于 ScFv 的筛选，可筛选出高亲和力和特异性更强的 ScFv。通过 C 末端连接另一单链抗体基因，可构建和表达出双特异单链抗体，应用于肿瘤的诊断和治疗。也可与 IL-2、IFN 等细胞因子基因连接，使细胞因子特异性杀伤肿瘤细胞，增强抗肿瘤活性，降低其毒副作用。随着有关研究的不断深入，ScFv 应用也会日趋广泛。

三、单 域 抗 体

只含 V 区基因片段的小分子抗体，就称为单域抗体，其相对分子质量仅为整个 Ig 分子的 1/12，故也称之为小抗体。

单域抗体只有 V_H 或 V_L 一个功能结构域制备相当简单；而且可以在大肠杆菌等宿主中表达出有活性的抗体片段。

单域抗体突出的特点是分子小，因此其抗原性很弱，较适合于体内的应用。由于其分子小，因而易于穿透组织，可用来快速清除组织内的毒素，并可渗透到完整 McAb 分子不能达到的部位，从而提高对肿瘤等的治疗效果。

与完整 McAb 相比，单域抗体的亲和力较低，而且相对的较有黏性。单域抗体最大的缺陷是不含 Fc 段，因此无固定补体及介导 ADCC 的作用。这将影响其实用价值。然而，通过不同的 V_H 基因以及其他功能分子的重组等，可能获得双特异性，甚至多特异性以及具有其他功能活性的单域抗体，但要进行这些改造还需很多的研究。

第四节　双特异性抗体

一、双特异性抗体的发展

天然 Ig 是由两个完全相同的 V_L 和 V_H 区域构成，该区域是特异性识别并结合抗原的关键部位。而经人工设计构建的双特异性抗体（bispecific antibody，BsAb）则由两个不同的抗原结合位点组成，可同时与两种不同的抗原决定簇结合，并可将其耦联的药物、酶或放射性核素等导向到靶部位，这种具有双价双特异性的抗体又称为双功能抗体（bifunctional antibody，BfAb）。

二、BsAb 的制备与构建

BsAb 的制备方法包括化学耦联法、杂交-杂交瘤法和基因工程方法。现介绍基因工程方法。

目前，一般采用与抗原结合的抗体片段，如 Fab、Fv 或 ScFv 为构件设计制备 BsAb。这种 BsAb 具有相对分子质量小，组织穿透力强，易于清除的特点，并由于其中缺乏恒定区，不会引起全抗体的生物学效应，减少了与表达 Fc 受体的细胞发生非特异性结合的机会。

制备 bisFv 的关键问题是如何使两个单一的蛋白融合成具有双功能性的蛋白质，这与天然蛋白质中两个或多个亚基装配成功能性聚集体相类似。目前，至少有 3 种比较系统的方案用来制备 bisFvs：①借助一种能提供蛋白质间相互作用的“二聚化结构域”，如双歧性的螺旋结构、亮氨酸拉链等，来充当抗体恒定区，从而实现单体向二聚体的转化，这种双价体称为“微抗体”（miniantibodies）。微抗体与完整抗体有相同的抗原特异性和相近的亲和力，其稳定性则高于单价的 ScFv。②以接头直接融合两种 ScFv。③将连接肽缩短（0～15 氨基

酸)，制备出名为“双体”(diabodies)的双价双特异性 ScFv。将 A 抗体的重链 V 区(V_HA)与 B 抗体的轻链 V 区(V_LB)通过上述短肽连接，构成杂合 ScFv，同理再以 V_LA 和 V_HB 构建另一杂合 ScFv。当二者同时表达时，由于短接头的限制，位于同一 ScFv 链的 V 区之间难以匹配，被迫与另一 ScFv 的、但实际是来源相同的 V 区相匹配，巧妙地形成有两个抗原结合位点的二聚体。

第五节　噬菌体抗体

噬菌体抗体库(phage antibody library)技术是通过 PCR 将全套人抗体重链和轻链 V 区基因克隆出来，并在噬菌体表面表达、分泌，经筛选后获得特异性抗体。所构建的抗体的集合称为抗体库(antibody repertoire)或组合抗体库(combinatorial antibody library)，从中筛选得到的抗体称为噬菌体抗体。

一、基本原理

该技术以噬菌体为载体，将抗体基因(*Fab* 或 *ScFv* 基因等)与噬菌体编码的外壳蛋白基因Ⅲ(*cp*Ⅲ)或 *V*Ⅲ(*cpV*Ⅲ)相连，在噬菌体表面以抗体-外壳蛋白融合蛋白的形式表达。经辅助病毒感染后，携带有表达载体的宿主菌就会释放出带有抗体片段的噬菌体颗粒。这种噬菌体抗体可以特异识别抗原，又能感染宿主菌进行再扩增，因此可以采用类似于亲和层析的原理从噬菌体抗体库中筛选出特异性抗体。

噬菌体抗体库技术的产生依赖于三项实验技术的发展：

1. PCR 技术的发展使得利用一组引物克隆出全套 Ig 可变区基因成为可能。目前扩增 V_H 和 V_L 基因的引物主要根据 FR 区、前导肽区或保守序列设计，利用这些引物可卓有成效地扩增多样化的可变区基因。为便于克隆，常在引物的两端分别设计合适的酶切位点。

2. 基因工程表达技术的发展与成熟使得在大肠杆菌中分泌有结合功能的 Ig 分子片段获得成功。抗体与抗原结合的部位是由两条肽链(V_H 和 V_L)通过复杂的空间折叠形成的立体结构，维持这一结构对抗体的功能至关重要。在 Ig 可变区基因 5′加入细菌的信号肽序列，Ig 可变区即可分泌到胞周间隙，并在那里完成立体折叠，成为有功能的蛋白质分子，还可形成链间二硫键，产生二聚体。这一成功说明了用抗体库在原核细胞筛选特异性抗体的可行性。

3. 有效的筛选系统(噬菌体表面呈现技术)的建立　将肽链通过与单链噬菌体外壳蛋白(pⅢ或 pVⅢ)融合表达在丝状噬菌体的表面，利用其可以再扩增的特性，将靶分子固定化，通过亲和筛选-洗脱-扩增，可筛选到靶分子的配体肽链或 Fab 段。

二、噬菌体抗体库的分类

按抗体基因的来源，噬菌体抗体库分为：

1. 天然抗体库(native antibody library)　以淋巴细胞为抗体 V_H 和 V_L 的基因来源所构

建的抗体库成为天然抗体库。

2. 免疫抗体库　这是相对于天然抗体库而言的，是指利用从特定抗原免疫的动物或人体获取的淋巴细胞为材料所构建的抗体库。此种抗体库的库容大多为 $10^6 \sim 10^8$，从这种抗体库中比较容易筛选出高亲和力的抗体。免疫抗体库对于筛选特定抗原的抗体是非常合适的。

3. 半合成抗体库　CDR 决定了抗原-抗体结合的特异性，其中重链 CDR1、CDR2 只能组成 7 种立体环状结构。而 CDR3 则由于有 30 种 D 片段、6 种 J 片段，以及 N 区的重排组合，显示出长度、序列、和空间结构的高度可变性，具有比其他任何一个 CDR 更重要的作用，是抗原结合的活性中心。这种 CDR 结构的特异性是半合成抗体库的基础。

根据 CDR3 的结构特点，人工合成长度为 5 个或 7 个氨基酸的 CDR3 随机序列，通过 PCR 法与胚系(germline)的 V_H 基因重组，利用这种形式获得的抗体基因所构成的抗体库称为半合成抗体库。

4. 全合成抗体库　抗体可变区基因中 CDR 序列完全是随机合成的，利用这种形式获得的抗体基因所构成的抗体库称为全合成抗体库。在全合成抗体库中，选择抗体可变区模板的依据是其在大肠杆菌中的表达能力，这样可以确保它们能呈现在噬菌体表面，而且可借助可溶性片段的形式大量生产。

三、噬菌体抗体的克隆和表达

(一) 抗体可变区基因获得

实验表明，杂交瘤细胞、免疫脾细胞、淋巴结细胞、外周血淋巴细胞，甚至体外免疫的细胞都可以作为克隆 B 细胞全套可变区基因的建库来源。有的研究提示，在这些细胞中，以选用淋巴结细胞建库为最好。因为这种细胞有利于筛选到高亲和力和特异性的抗体。

目前，获得可变区基因的途径主要有两种，一是从基因文库中获取可变区基因，二是用 PCR 法扩增 Ig 的 V_H 和 V_L 基因。其中，PCR 是目前在构建噬菌体抗体库时最常用的一种方法。模板的完整性及其浓度的高低，引物的设计，退火温度和时间的选择，循环次数的多少等，均可能影响 PCR 的结果，其中引物的设计尤为关键。

(二) 表达载体

噬菌体抗体库的表达载体可分为 λ 噬菌体(phage λ)、单链丝状噬菌体(filamentous phage)及噬菌粒(phagemid)三种系统，这三种系统各有其利弊。λ 噬菌体表达系统可通过菌斑印迹(plaque lift)筛选克隆，但外源性片段不能太大，其库容较小(10^6)。丝状噬菌体主要为 Winter 实验室所采用，目前已构建了 fd CAT1 及 fd DOG1，并已制备 ScFv；此载体中又包括单价及多价两个子系统，可筛选出高亲和力抗体。噬菌粒是应用最多的表达载体，也是简便高效模拟 B 细胞产生抗体的原核表达系统。目前对此系统已改建的主要表达载体包括：以 pBluescript 构建的 pComb3、pComb8、PCBAK8；以 pUC 构建的 pHEN1 及 pCANTAB；以 pDS31-1 构建的 PSEX；以 pUC、PGC1 构建的 pTac cp；以 pSW1 构建的 pEXmide 3 等。

这些表达系统在表达抗体时，一般都较质粒表达系统的产量高、活性强、特异性好，易于筛选，其本身也可与质粒联合应用，易于构建抗体可变区基因文库。其中的丝状噬菌体及噬菌粒还可作为寡核苷酸定位诱变的中间载体，同时具有诱变及表达功能，这为模拟体细胞的突变提供了前所未有的工具。

（三）抗体库的构建

在构建表达 Fab 段的抗体库时，可把 Fab 的一条链（Fd）在噬菌体表面呈现，另一条链（κ 或 λ 轻链）克隆到 pASK22 类型的载体内，二者转染同一大肠杆菌，分泌到胞周间隙中的一条链与噬菌体表面的另一条链相匹配并联合成为在噬菌体表面呈现的功能性 Fab 抗体。

在构建表达 ScFv 的抗体库时，其方法有多种，如将 V_H 和 V_L 基因分别克隆到不同噬菌体载体上再将两载体在体内或体外重组，或将 V_H 和 V_L 依次克隆到同一噬菌体载体上，也可用细胞内 PCR 装配技术在淋巴细胞内将 V_H 和 V_L 基因连接，再克隆到噬菌体载体上，还可将 V_H 基因库与单一 V_L 基因组合等等。此外，在抗体基因与 *cp*3 之间插入一个琥珀（amber）中止密码子，当此密码受抑制时抗体则分泌表达。因此当重组噬菌体生长在抑制型或非抑制型大肠杆菌时就可分别模拟 B 细胞和浆细胞功能。

（四）转化大肠杆菌

噬菌体抗体库的大小与大肠杆菌转化率明显相关。400μl 感受态的大肠杆菌经电穿孔转化，至少有 10^8 个细菌可被噬菌粒载体转染，这与体内抗体库独特型数目一致。如果噬菌粒在体外包装，则还可以明显增加转染大肠杆菌的数量，因此电穿孔是一种理想而有效的增加大肠杆菌抗体库数量的手段。

（五）噬菌体抗体的筛选和富集

从噬菌体抗体库中筛选表达特异性抗体的经典方法主要有两种：①将纯抗原包被在固相介质上，如酶标板、免疫试管或亲和层析柱，然后加入待筛选的噬菌体，洗去非亲和性或低亲和性结合的噬菌体，回收高亲和性的噬菌体；②将抗原与生物素基团相连，再将其固定在包被有链霉素抗生物素（streptavidin）的磁珠上对噬菌体进行筛选。

四、噬菌体抗体的应用

近年来，噬菌体抗体已成为基因工程抗体研究的热点之一，并被广泛应用于生命科学的各个领域。

（一）基础研究

许多自身免疫性疾病的发生是由于机体产生的高亲和力自身抗体借助 ADCC 作用和（或）激活补体系统等机制引起的组织病变。目前，用于研究自身抗体的方法有杂交瘤技术、EBV 转化 B 细胞和噬菌体抗体库技术。其中，由于种属间的差异限制了鼠单抗的应用价值，而人杂交瘤和 EBV 转化 B 细胞技术制备的单抗多为 IgM（疾病相关的器官特异性血清

中自身抗体主要为 IgG)。相比之下,噬菌体抗体库技术易于操作,而且能够同时获得大量自身抗体。据统计,1993～1995 年间文献报道的 IgG 类自身抗体有 64 种,其中 49 种是通过噬菌体抗体库技术得到的。1996 年,McIntosh 等人从抗体库中一次性得到 19 种抗甲状球蛋白的人单抗。近年来,以噬菌体抗体库技术为工具对各种自身免疫性疾病的研究报道有:器官特异性自身免疫性疾病(桥本甲状腺、Grave 病、溃疡型结肠炎、舍格伦综合征、重症肌无力、原发性胆汁硬化症)以及系统性自身免疫性疾病(系统性红斑狼疮、魏格纳肉芽肿)。可以预见,利用噬菌体抗体有可能搞清楚自身抗体以及 B 细胞在提呈自身抗原方面的作用机制,最终可能借助 X 射线晶体学解析出自身抗体所识别的抗原表达。

(二) 诊断和治疗

噬菌体抗体库技术与单抗技术相比,其优势在于可不经免疫制备抗体,能够制备人单抗,从而避免了鼠单抗应用于人体时发生的免疫反应。而且,通过体外亲和力成熟技术和动力学常数筛选技术,能够对有关抗体进行改造,从中得到特异性和动力学特性均较好的抗体。例如,体内影像诊断用的抗体应具有高亲和力、快的解离率及短的半衰期,而靶向治疗抗体则需要慢的解离率和长的半衰期。噬菌体抗体库技术则可根据需要把有关的抗体改造成诊断试剂或者治疗药物。

利用噬菌体抗体库技术已经获得了大量针对病原微生物的抗体,如 HAV、HBV、HCV、HIV、VZV (varicella zoster virus)、RSV (Rous sarcoma virus)、风疹病毒、CMV (cytomegalovirus)、HSV、流感嗜血杆菌和破伤风毒素等。这些噬菌体抗体中,许多具有临床诊断和治疗的应用前景。

五、展　　望

噬菌体抗体库技术模拟生物体内 B 细胞的有关特性——识别与扩增的统一,是一种极为高效的表达、筛选抗体的体系。而且,该技术使不经免疫而直接利用抗原从抗体库中筛选特异性的抗体变为现实,可绕过免疫制备全人源性抗体,避免了鼠源性抗体在人体应用时诱发的 HAMA 等不良反应。解决了人杂交瘤系统低效性的难题,使人单克隆抗体的制备有了突破。这是基因工程技术的重大发展。因此,噬菌体抗体库技术的出现及噬菌体抗体的研制成功已成为生命科学研究的突破性进展之一,必将为生物技术领域的发展带来巨大的推动。

但是,目前的抗体库技术仍有许多问题尚需进一步解决。首先,以丝状噬菌体表面呈现系统作为筛选手段的抗体库技术,其前提是文库中的抗体片段在大肠杆菌中必须获得分泌性表达。已有的研究表明,并非所有的抗体片段都能在大肠杆菌中进行分泌性表达。如何将抗体库储存的信息通过原核系统完整地再现出来,尚需进一步的探讨。其次,由于道德、伦理方面的制约,不能免疫人而产生较高特异性的人单抗。虽然通过构建足够多样性的天然抗体库,同时结合体外亲和力成熟等方法可以产生较高亲和力的抗体,但其库容总量、多态性及筛选等技术环节尚待深入研究。再则,这种基因工程抗体都是在 *E.coli* 中表达的 Fab 片段或单链抗体,其表达量都很低,难以用于实践。因此,建立高效表达系统有助于此

问题的解决。

随着该项技术的发展和完善，人们可以在更大的程度上根据需要而改造和制备有关的抗体。可以预见，一个随意和定向地制造抗体的时代已为期不远。

（马春红　韩丽辉）

第十四章　真核生物基因表达调控

Chapter14. Regulation of Eukaryotic Gene Expression

第一节　真核生物基因表达调控的基本理论

一、真核生物基因的结构功能特点

真核生物在基因结构上与原核生物有相当大的差异,其特点如下:

(一)真核生物的DNA含量大,具有复杂的高级结构

真核生物的DNA含量约为大肠杆菌DNA含量的数百至数千倍,与裸露状态的原核生物DNA不同,真核生物DNA大多与蛋白质结合,形成复杂而又有序的高级结构即染色质或染色体,并由核膜包裹起来,形成细胞核(线粒体与叶绿体DNA除外)。这就决定了真核生物基因表达必然受到其高级结构的影响,并且转录与翻译过程不能同步进行,从而可以在更多的水平上进行调节控制。

(二)真核生物的编码基因大多是不连续的

绝大多数真核生物(特别是高等脊椎动物)的基因是由外显子与内含子镶嵌排列而成的。所谓外显子(exon)是指为蛋白质氨基酸序列编码的那部分DNA序列;而内含子(intro)则是指那些插入到编码序列之中的非编码序列。基因转录产生的初级转录产物必须经过一定的加工过程,将插入的内含子序列切除,才能形成成熟的mRNA分子。这些内含子的功能现在仍不清楚,据推测可能有如下几种作用:①某些内含子中含有基因表达的调控序列,参与表达的调控;②不同的剪切方式可产生不同的基因表达产物;③提供遗传变异机会,与生物进化有关。

(三)真核生物的DNA中含有大量重复序列

所谓重复序列是指在基因组中多次反复出现的DNA序列。这些重复序列按其出现的频率可分为低度、中度和高度重复序列。各种重复序列的长度、序列及重复次数不同,其功能亦有极大差别。有些重复序列是为蛋白质编码的,有些则不是编码序列。其功能可能大致有以下几种:

1. 某些基因产物的生物学功能极为重要,且需要量极大,如组蛋白、rRNA、tRNA,及某些糖代谢酶类等,因此,为这些分子编码的基因以多拷贝形式存在,可充分满足细胞的需要

量。这也是表达调控的一种有效方式。

2. 某些重复序列与染色质构象、着丝点的形成有关。

3. 参与基因复制及表达的调控:某些高度重复序列中 AT 含量丰富,易解链,有利于与调控基因复制和表达的蛋白质因子结合;另外,许多重复序列结构中含有末端反向重复序列,可形成茎环状结构,这种高级结构参与基因表达调控,也可能与基因的转位有关。

(四) 功能性基因和假基因

某些有一定同源性而又不完全相同的基因簇,组成一个基因家族,如珠蛋白基因、*ras* 基因家族等。它们分工行使不同的生物学功能,称为功能性基因。还有一些则根本不具有功能,称为假基因(pseudogene),它们可能只是进化的痕迹,也可能为今后的进化提供物质基础。例如,多种蛋白质水解酶都具有相似的活性中心,可以推断,它们都是由同一种原始蛋白水解酶进化而来的。这种原始酶基因通过扩增形成多拷贝基因,然后通过突变、自然选择而进化为各种不同功能的蛋白水解酶。

(五) 真核生物基因组中无"超基因"式的操纵子结构

真核生物基因组中不存在像原核细胞"超基因"式的操纵子结构,功能相关的基因大多分散在不同的染色体上,即使位置相近,也是分别进行转录的,而不产生多顺反子 mRNA。

(六) 真核生物基因组与生长有关基因的调控相对较为简单,而与分化有关基因的调控则较为复杂

原核生物生存的惟一目的就是无限生长繁殖,因此其基因表达调控系统的作用就是在一个特定环境中为细胞提供最大的生长速度,即将酶活性控制在最大速度所要求的点上,并使细胞尽快地适应变化着的环境。而真核生物,特别是多细胞高等生物,则除了生长繁殖外,更重要的是要进行分化。由于高等动物的血液循环系统为细胞提供了一个较稳定的环境,因此生长的调控相对较为简单,而分化的调控则复杂得多。高度分化的组织细胞中,只有大约 10% 的基因不同程度地表达,大多数基因被关闭,如何调节各种基因的表达活性,其复杂程度是可想而知的。而原核生物中一般有 40%～50% 的基因处于活性状态。

二、真核生物基因表达调控的策略

如上所述,真核生物基因的表达是一个多阶段过程,因此,其表达的调控也是在多阶段水平来实现的,即转录前、转录、转录后、翻译和翻译后等五个水平。

(一) 转录前的表达调控

转录前的调控是指发生在基因组水平上的基因结构的改变,包括基因丢失、基因扩增、基因重排、甲基化修饰及染色质结构改变等。这种调控方式较稳定持久,甚至有些是不可逆的,主要见于机体发育过程中的体细胞分化。显然,这种调控方式未免缺乏灵活性、不适于对生长因子等刺激的应答反应。

1. 基因丢失　体细胞分化过程中，必须将某些基因永久性关闭。显然达到此目的的最简单有效的方式就是将这些基因丢失掉。在一些低等真核生物（如线虫、原生动物和昆虫等）的体细胞发育过程中确实发现了染色体丢失现象。如原生动物尖毛虫细胞中有大、小两个细胞核。小核中 DNA 是完整的，不具有转录活性，起维持种系的作用；而大核中只保留了部分 DNA，但具有转录活性。这种基因丢失现象并非普遍存在。

2. 基因扩增　基因扩增或基因放大（gene amplification）是指基因拷贝数的增加，是基因表达活性调节的有效方式之一。当发育分化或环境条件的改变使机体对某种基因产生的需要量剧增，而单纯靠调节其表达活性不足以满足需要时，则需要增加这种基因的拷贝数。扩增的基因可游离存在，也可整合入染色体，有些生物中可形成多条染色体。值得注意的是，不适当的基因扩增可导致某些疾病的发生。如某些癌基因（c-*myc*、c-ki-*ras*）的扩增可能与某些肿瘤的发生有关。

3. 基因重排　基因重排是指某些基因片段改变原来存在的顺序而重新排列组合。重排可以仅仅是空间位置或方向的不同，也可同时伴有某些基因片段的扩增或丢失。基因重排方式是基因表达的转录前调控的较重要方式之一。广义的基因重排有多种方式：①类似于 mRNA 加工的"剪切"方式；②转座子方式；③染色体转位方式；④外源基因序列（如反转录病毒长末端重复序列，LTR）的插入激活方式等。

(1) "剪切"方式：哺乳动物中免疫球蛋白（Ig）的产生是此种方式的代表。机体可产生上百万种不同的 Ig，Ig 的这种多样性是通过基因重排实现的。虽然编码 Ig 的基因数量不多（不足 500 个），却可通过不同的重排方式产生多种多样的抗体。κ 基因通过重排产生 $4\times200=800$种不同的轻链，而 *H* 基因可产生 $200\times10\times4=8000$ 种不同的 *VDJ* 连接可能性。这样，不考虑恒定区的不同，总共有 $800\times8000=6\ 400\ 000$ 种不同抗体。这种重排方式是极为经济的，节省了大量的基因组空间。

(2) 转座子方式：转座子（Tn）是基因组上可移动的遗传因子。当转座子插入到染色体上的某些地方时，就可控制此基因的表达，这种基因转移方式称转座子方式。转座子将一些原来在染色体上相距甚远的基因连接到一起，从而产生一些新的蛋白质分子或新的表达调控方式。如图 14-1 所示，Tn 由 5 个部分组成：①反向末端重复序列（IR）：这是转座酶识别位点，通过酶切、复制后再插入到其他位置；②转座酶基因：编码转座酶，转座酶可识别插入顺序，并在 DNA 上切割；③阻遏蛋白基因：编码阻遏蛋白，对转座起负调控作用；④调控区：为调控蛋白的结合位点；⑤附加基因：与转座无关，如 Tn3 带有青霉素抗性基因——β-内酰胺酶基因。转座子在进化中的重要作用是无疑的，但其对于真核生物基因表达的调控作用，仍有待进一步研究。

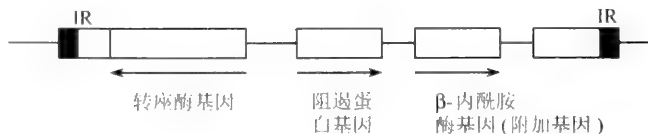


图 14-1　转座子 Tn3 结构示意图

(3) 染色体转位方式(chromosome translocation):严格来说,染色体转位方式并不是基因表达的正常调控方式,而是一种病理现象,但确实可以导致某些基因的表达改变。此方式常见于一些癌基因的活化过程。如人 Burkitt 淋巴瘤中常可见 t(8;14)染色体易位,导致原来不具转录活性的 c-*myc* 癌基因转位到 Ig 重链基因的强大增强子作用之下,从而被激活转录。

(4) 外源基因插入激活方式:类似于转座子方式,但可能并无转座子的参与。主要见于反转录病毒的长末端重复序列(long terminal repeat sequence,LTR)的插入引起的细胞癌基因的激活。LTR 可能具有增强子作用。也有人认为有一些(也可能是全部)RNA 肿瘤病毒本身就是转座子。

4. 甲基化修饰 某些高等生物,尤其是脊椎动物中,DNA 上特定的 CpG 序列(转录起始区)处的胞嘧啶可发生甲基化修饰(5mC),这种甲基化修饰可以阻止某些基因的转录,并且能遗传到子细胞中去。研究表明,转录活跃的基因是低甲基化或不甲基化的,而不表达基因则高度甲基化。如珠蛋白基因在红系细胞中是低甲基化的,而在不表达珠蛋白的细胞中则高度甲基化。持续表达的看家基因(house-keeping gene)的转录起始区极少发生甲基化(所谓看家基因是那些维持细胞正常功能所必需的、持续表达的基因)。有人认为 X 染色体的失活也是 DNA 甲基化的结果。在脊椎动物中,DNA 甲基化是普遍现象,但在无脊椎动物中则较少,而在昆虫中则根本没有。

5. 染色质结构对基因表达的调控作用 真核生物基因组的最重要特征之一就是 DNA 与组蛋白、非组蛋白等多种蛋白质和少量 RNA 及其他物质结合,形成染色质(chromatin)或染色体(chromosome)结构。

含大量碱性氨基酸(约 25%)的组蛋白与 DNA 紧密结合,可保护 DNA 免受损伤,维持基因组的稳定性,并抑制基因的表达。去除组蛋白则基因转录活性增高。这种组蛋白的结合与解离是真核生物基因表达调控的重要机制之一。机体可通过磷酸化、乙酰化、甲基化等机制将碱性组蛋白修饰,从而减少正电荷,减弱与 DNA 的结合能力。

染色质有结构紧密的超卷曲状态和结构松散的伸展状态及多种中间状态。这两种状态的转换主要与组蛋白 H1 有关,H1 与伸展状态的核小体结合形成螺旋管结构。结构紧密的染色质为异染色质;结构松散的染色质为常染色质。转录活性较高的基因都位于结构较松散的常染色质中,而位于结构紧密的异染色质中的基因则大多不具转录活性。

转录活跃的基因位于异染色质区域,由于结构松散,去除了组蛋白的保护作用,因此对于核酸内切酶Ⅰ(DNaseⅠ)水解作用较敏感,称为 DNaseⅠ敏感区。而结构紧密的常活性染色质则对 DNaseⅠ不敏感。因此常将 DNaseⅠ敏感性作为该基因的转录活性的标志。另外,在活性染色质中还存在一些对 DNaseⅠ特别敏感的区域,称为 DNaseⅠ超敏感(DNAase Ⅰ hypersensitive site)。DNaseⅠ超敏感区一般位于活性基因的 5′端,可能反映了基因转录的起始位点。

(二) 转录水平的调控

转录水平的调控是真核生物基因表达调控中最重要的一步,主要涉及以下三种因素的相互作用:一是 RNA 聚合酶(RNA ploymerase,RNA pol)的调节;二是顺式调控元件(*cis-*

acting element);第三是反式作用因子(*trans*-acting factor)。

1. RNA 聚合酶　与原核生物单一的 RNA 聚合酶不同,真核生物的 RNA 聚合酶有三种:RNA 聚合酶Ⅰ、Ⅱ、Ⅲ。RNA 聚合酶Ⅰ存在于核仁中,其转录产生 r RNA(5.8S、18S 和 28S r RNA),对 α-鹅蕈碱有一定的抗性。RNA 聚合酶Ⅱ存在于核质中,其转录产物为 mRNA 及其他一些功能不明的相对分子质量小的 RNA,如 $U_{1\sim6}$ RNA,对 α-鹅蕈碱十分敏感,与细胞生长、分化直接相关,其表达调控最为复杂。RNA 聚合酶Ⅲ存在于核质中,其转录产物为 5S rRNA 和 tRNA;对 α-鹅蕈碱的敏感性介于上述两者之间。

2. 顺式调控元件　是与结构基因串联、对基因转录的精确起始和活性起重要调节作用的特定 DNA 序列。真核生物中,不同的顺式调控元件按不同的数量、类别及空间位置串联排列,组成了各基因表达的调控区域,它们的协同作用决定了基因转录的精确起始与转录效率。

(1) 启动子(promoter):启动子是与启动基因转录有关的核酸序列,位于基因转录起始位点 5′端,只能在近距离起作用(一般在 100bp 之内),有方向性,空间位置较恒定。它是转录的最基本信号结构(详见第八章)。

(2) 增强子(enhancer):增强子是一类能促进基因转录活性的顺式调控元件,但它本身不具备启动子的活性。其特点是:①无方向性;②远距离作用,距靶基因可近可远,甚至远至几十个 kb 也同样能发挥作用,可位于基因的上游、下游或内部;③无基因特异性,对各种基因启动子均有增强作用;④具有组织特异性;⑤具有相位性,它的作用虽然与距离无关,但只有当它位于 DNA 双螺旋的某一相位时,才具有较强活性。典型的增强子是 SV40 早期基因中的 72bp 重复序列,其中有由特定核苷酸组成的核心序列(core sequence,module,motif)。

(3) 加尾信号及转录终止信号:在加 polyA 尾位点的上游 10～20bp 处,常见一保守的 AATAA 序列,如去除此序列,基因会连续转录下去而不终止,被认为是加尾信号,但与转录终止的关系仍不能肯定。

3. 反式作用因子　反式作用因子(又称为反式作用转录因子)是由位于不同染色体或同一染色体相距较远的基因编码的蛋白质因子,通过与顺式调控元件和 RNA 聚合酶的相互作用而调节基因转录的活性(详见第八章)。

(三)转录后水平的调控

从原始 mRNA 到成熟 mRNA 的调节,是表达调控的一个重要内容。

1. 戴"帽"(capping)　mRNA 转录后不久即在其 5′末端加上 m^7GpppN 的"帽子"。其作用是:①防止 mRNA 的降解,延长寿命;②与核糖体小亚基结合;③被翻译起始因子所识别。

2. 加"尾"(tailing)　除组蛋白 mRNA 外,真核生物 mRNA 的 3′末端均有一 poly A 序列。加尾信号是 AAUAAA,由核酸内切酶切开 RNA3′末端(常在 GC 序列之后),然后加上 polyA。polyA 的功能是:①是 mRNA 进入细胞质所必需;②保持 mRNA 稳定性,延长寿命。

3. 甲基化修饰　主要是形成 6-甲基腺嘌呤(6mA),其意义不明。

4. 拼接(splicing)　将 hnRNA(不均一核 RNA)中的内含子序列切除,外显子部分连接起来,称为 RNA 拼接。在拼接位点上常有一些保守序列,在其 3′端多为 AG,在其 5′端多为 GU,它们是指导正确拼接的信号。拼接过程有核酸内切酶的参与,有些 RNA 则可自身催化自发拼接,一些小分子 RNA,如 snRNA,也参与 mRNA 的拼接过程。

大多数基因的 mRNA 加工较简单，只产生一种成熟的 mRNA，指导合成一种蛋白质分子。但有些基因的 mRNA 可有不同的加工方式，产生不同的蛋白质。这对于细胞的功能分化有一定的作用。如某些基因有多个加 polyA 位点，通过不同的拼接方式产生两种不同的蛋白质，如 IgM、D、E、G 重链基因、大鼠降钙素基因、人血纤维蛋白原基因等。Igμ 链基因有两个 poly A 位点，通过不同的拼接方式产生两种不同的蛋白质，较短的 μs 型蛋白质是分泌型可溶性蛋白，而较大的 μm 型蛋白则结合于细胞膜上。

（四）翻译水平的调控

翻译过程主要涉及 mRNA、tRNA、核糖体和可溶性蛋白因子四大类装置。对于翻译水平调控的机制目前了解不多，可能有以下几个方面：

1. 对可溶性蛋白质因子的修饰　肽链起始因子 eIF-2 在激酶的作用下磷酶化后，可抑制蛋白质的合成。这可能是因为磷酸化的 eIF-2 与 eIF-2B 结合后不能解离，从而影响 eIF-2 的再循环利用。

2. 对 mRNA 稳定性的调节　虽然机制还不清楚，但细胞确实可以调节各种不同 mRNA 的寿命。起重要作用的 mRNA 的寿命比其他 mRNA 的寿命长。如蚕蛹羽化成蛾后，需要大量的蛋白水解酶溶解蚕丝蛋白，此时蚕丝蛋白水解酶 mRNA 的半衰期长达 100 小时，而其他的 mRNA 半衰期只有 2.5 小时。催乳激素可使乳腺组织中的酪蛋白 mRNA 的半衰期提高 20 倍。

3. 反义 RNA 对翻译的调控作用　所谓反义 RNA(anti-sense RNA)是一段与 mRNA 互补的碱基序列。因此，反义 RNA 通过碱基配对与相应的 mRNA 结合，形成的双链 RNA 影响 mRNA 的模板效率，抑制翻译。此外，反义 RNA 也可能会抑制基因的转录，有些反义 RNA 可能参与抑制 mRNA 的拼接过程。

天然反义 RNA 首先在原核生物中发现，如 Tn10 编码的转座酶基因就受到反义 RNA 的调控，使得转座速度不会受到转座酶基因拷贝数的影响。近年来，在真核生物中也发现了一些可能的天然反义 RNA。如禽类 *c-myc* 基因可反向不连续转录出 3 个反义 RNA，分别位于第 1 个外显子上游、第 1 个内含子及第 2 个内含子和第 3 个外显子。推测它与 mRNA 拼接、翻译及转录调控有关。

总之，在翻译过程中，确实存在一系列的调控机制，有些是非特异性的，有些是特异性的针对某特定基因的。这方面仍有很多的课题有待进一步研究。

（五）翻译后水平的调控

蛋白质合成后，还须经一系列的加工过程才能成为有活性的功能蛋白质。

1. 切割　分泌型蛋白质必须切除疏水性信号肽。多种酶是以酶原形式存在的，如胰蛋白酶原，必须经酶解才能成为活性蛋白。

2. 化学修饰　①磷酸化：磷酸化-去磷化作用是机体调节酶活性的最常见方式。磷酸化位点有丝氨酸、苏氨酸和酪氨酸；②糖基化：糖基化位点多为天冬氨酸、丝氨酸和苏氨酸；③乙酰化：主要是 N 末端及赖氨酸。

3. 连接　有些蛋白质经切割成小肽后，还必须以一定的方式连接起来。如红细胞凝集素、外源凝集素及刀豆球蛋白等。因此，推测细胞中可能存在一类多肽连接酶。

第二节 基因表达调控的研究方法

基因表达调控研究已成为当代分子生物学中的重要课题。技术的进步促进了调控研究的逐步深入,而研究的深入又产生了对新技术的需求,因而新方法新技术层出不穷,种类繁多,本书仅择其重要者进行介绍。

一、转录前表达调控的研究方法

(一) DNase Ⅰ超敏感分析(DNase Ⅰ hyperseneitivity assay)

活跃转录的基因存在于结构松散的常染色质上,较少受到蛋白质的保护,因而易受核酸内切酶 DNase Ⅰ 的攻击,特别是基因转录调控位点。这种对 DNase Ⅰ 的高度敏感性即称 DNase Ⅰ 超敏感性。

此方法基本步骤是分离细胞核,与一定量的 DNase Ⅰ 反应后提取 DNA,电泳分离后,进行 Southern 杂交分析,根据片段的大小,鉴定超敏感位点。DNase Ⅰ 超敏感区一般位于活性基因的 5′端,可能反映了基因转录的起始位点。

(二) DNA 甲基化分析

在哺乳动物细胞 DNA 中,CpG 核苷酸序列出现的频率较低,只有预计频率的 1/5 左右,尽管只有一部分 CpG 核苷酸被甲基化,但这种甲基化具有高度的细胞特异性。DNA 上特定的 CpG 序列处的胞嘧啶甲基化修饰对基因的表达有抑制作用。检测 DNA 的甲基化状态有助于了解该基因的转录活性。甲基化分析的方法有多种,最简单有效的方法是利用限制性内切酶对甲基化的敏感性不同进行测定。对 CpG 序列甲基化作用敏感性不同的限制性内切酶见表 14-1。

表 14-1 对 CpG 序列甲基化作用敏感性不同的限制性内切酶

限制性内切酶	可切割序列	不可切割序列
Hpa Ⅱ	CCGG	C_mCGG
Msp Ⅰ	CCGG, C_mCGG	$_mCCGG$ ·
Cfo Ⅰ, *Hha* Ⅰ	GCGC	G_mCGC, G_mCG_mC
Sma Ⅰ	CCCGGG	CC_mCGGG
Xma Ⅰ	CCCGGG	$_mCCCGGG$
	CC_mCGGG	
Acc Ⅱ, *Bsu* E Ⅰ	CGCG	$_mCGCG$
Bstu Ⅰ	CGCG	$_mCG_mCG$
Fnuu Ⅱ	CGCG	$_mCGCG$, CG_mCG_s
Tha Ⅰ	CGCG	$_mCGCG$, $_mCG_mCG$
Xho Ⅰ	CTCGAG	CT_mCGAG
Ava Ⅰ	CP_yCGP_uG	$CP_{ym}CGP_uG$
Sal Ⅰ	GTCGAC	GT_mCGAC
Taq Ⅰ	TCGA, T_mCGA	TCG_mA
Aos Ⅱ, *Aha* Ⅱ	GP_uCGP_yC	$GP_{um}CGP_yC$

续表

限制性内切酶	可切割序列	不可切割序列
Cla Ⅰ	ATCGAT	AT_mCGAT
Hae Ⅱ，*Ngo* Ⅰ	P_uGCGCP_y	$P_uG_mCGCP_y$
Nar Ⅰ	GGCGCC	GG_mCGCC
	$GGCGC_mC$	
Mlu Ⅰ	ACGCGT	A_mCGCGT
Mae Ⅱ	ACGT	A_mCGT

二、转录水平调控的研究方法

转录水平调控的研究方法很多，包括转录起始位点的测定、进行中的核转录分析、不同细胞间表达水平的差异文库分析、调控元件研究的报告基因法。

(一) 转录起始位点的测定

精确测定基因的转录起始位点，对于进一步研究其基因表达调控机制具有极重要的作用。常用的转录起始位点测定方法有三种：核酸酶 S1 作图法(S1-mapping)、RNase 作图法(RNase mapping；或 RNase 抗性分析法，RN-resistance assay)和引物延伸法等。

S1 及 RNase 作图法的基本原理如图 14-2 所示，5′末端标记的单链 DNA 或 RNA 探针可与相应的 mRNA 互补链杂交，用核酸酶 S1 或 RNase 消化未杂交的单链部分，测序胶电泳分离，放射自显影测定 DNA 片段的长度，估计 5′标记末端与 mRNA 5′末端的距离，即可判断出转录起始位点。

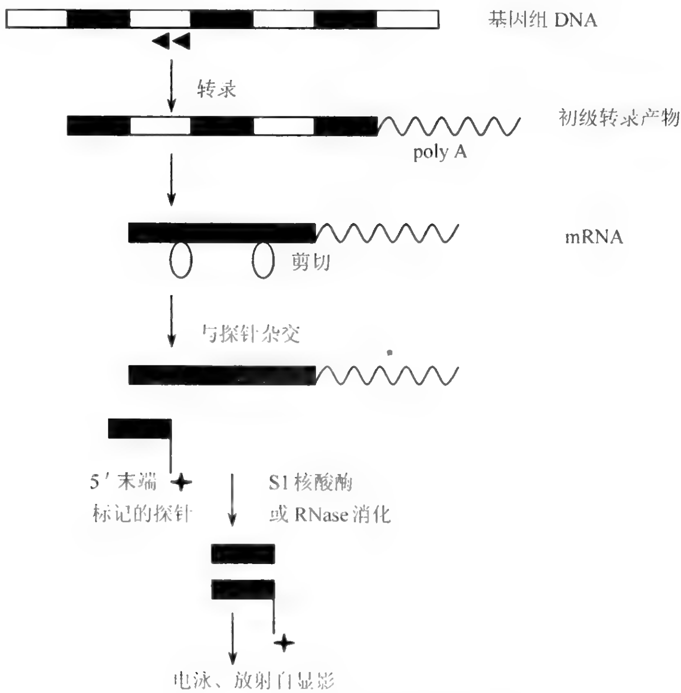


图 14-2　S1(RNase)作图法原理示意图

引物延伸法的原理是：待检基因 mRNA 与过量 5′末端标记的单链 DNA 引物杂交，利用反转录酶进行延伸反应合成 cDNA，然后，测定合成的 cDNA 产物的长度，即可判断出引物 5′末端与 mRNA 5′末端（即转录起始位点）的距离。

（二）进行中的核转录分析

进行中的核转录分析（nuclear run on transcription），即在细胞核提取物中加入放射性核素标记的核苷三磷酸，使其掺入到正在转录的 mRNA 分子中。通过核酸分子杂交方法，即可鉴定出多种基因是否转录及其转录的量。这是测定基因转录活性的较为可靠的方法，排除了 Northern 杂交等方法中 mRNA 半衰期对实验结果的影响。可同时分析多种基因的转录活性，也是此法的一大优点。

（三）不同组织细胞或同一细胞不同状态基因转录的差异分析（差示文库）

差示文库（differential library）又称扣除文库（subtracted library），是反映不同组织细胞或同一种细胞处于不同功能状态下的基因表达差别的 c DNA 文库。这对于研究细胞发育分化、细胞分裂周期、细胞对药物和生长因子的诱导反应及肿瘤等疾病的分子基础都有极大的帮助。

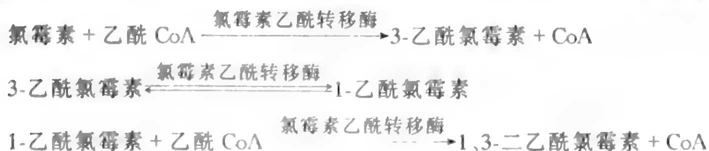
建立差示文库的基本流程是：分别提取不同细胞的 mRNA，将其中一种 mRNA 反转录成 cDNA，然后与过量的另一种细胞 mRNA 进行杂交，羟基磷灰石柱层析分离纯化未形成杂交体的单链 cDNA，它反映此细胞中特异性表达的基因。以所得到的单链 cDNA 为模板，合成第二链，以此双链 cDNA 建立 cDNA 文库，即为差示文库。

（四）启动子活性的研究（“报告”基因方法）

对假定顺式调控序列的启动子活性的测定，常采用“报告”基因（reporter gene）方法。即通过基因重组方法将待测 DNA 序列与一特异的“报告”基因相连，然后导入细胞中进行表达。通过测定表达出的“报告”基因产物的量，即可推断出此 DNA 序列是否具有启动子活性。报告基因为鉴定和分析基因的调控元件提供了便利的方法。

作为“报告”基因必须满足下列条件：①此基因所编码的酶活性必须易于同受体细胞内相似的酶活性相区别；②不会受细胞内其他酶活性的干扰；③测定方法简单、快速、灵敏度高、特异性高。目前有三种基因较为理想：氯霉素乙酰转移酶（CAT）基因、β-半乳糖苷酶基因和荧光素酶基因。

1. 氯霉素乙酰转移酶基因　氯霉素乙酰转移酶（chloramphenicol acetyltransferase，CAT）基因最初来自细菌转座子 Tn9，其产物 CAT 可产生氯霉素抗性。其催化的反应如下：



因此，可用^{14}C标记的氯霉素或^{14}C标记的乙酰辅酶A方便地进行CAT活性的检测。

2. β-半乳糖苷酶基因　β-半乳糖苷酶可作用于底物ONPG(邻硝基苯-β-D-半乳吡喃糖苷)或X-gal(5-溴-4-氯-3-吲哚-β-D-半乳糖)产生淡黄色或蓝色反应产物。β-半乳糖苷酶基因常作为CAT活性测定的内对照基因，以排除转染及蛋白质提取等过程中的误差对实验结果的干扰。

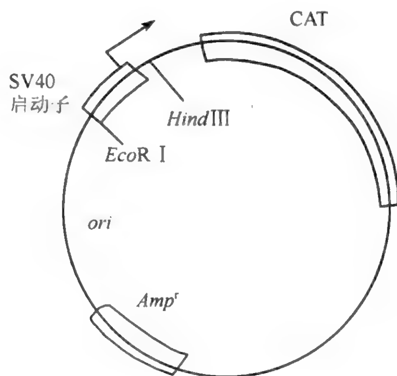


图14-3　pSVoCAT载体

如图14-3所示：切除SV40启动子，将外源启动子DNA片段插入，在大肠杆菌中扩增。然后与阳性对照(含SV40启动子)、阴性对照(不含启动子)分别导入哺乳动物细胞中进行瞬时表达。可同时将β-半乳糖苷酶基因共转染入细胞中，作为CAT测定的内标。提取全细胞蛋白提取物，与氯霉素和乙酰CoA一起保温，用薄层层析等方法测定乙酰化的氯霉素的量，即可判断DNA序列的启动子活性。

3. 荧光素酶基因　荧光素酶基因是近几年发现的一种新方法，是用于哺乳动物细胞的非放射性核素遗传报告系统。荧光素酶催化的生物发光反应需要荧光素作底物，以及ATP、Mg^{2+}及分子氧。将这些试剂与含有荧光素酶的细胞裂解液混合即会产生一种迅速衰减(在1秒内)的闪光。这种光信号可用一种配备了便于迅速混合反应物的自动注射装置的荧光检测仪进行检测。发光量的总值与样品的荧光素酶活性成正比，因而可对荧光素酶报告基因的转录进行间接估计。荧光素酶易被蛋白酶降解，因而在转染的哺乳动物细胞中的半衰期约为3h。荧光素酶的这种迅速更新的性质，使之成为研究可诱导系统的一种理想的候选报告蛋白。

荧光素酶作为报告基因具有很多优点：①荧光素酶的测定比氯霉素酰基转移酶(chloramphenicol acetyltransferase，CAT)或其他常用的报告基因更为灵敏。荧光素酶分析独有的高灵敏度使它可用于分析弱启动子及允许在转染实验中使用少量的DNA和细胞。②测定荧光素酶不需要用放射性物质。③荧光素酶在哺乳细胞中的半衰期为3小时，在植物中的半衰期为3.5小时。由于半衰期短，故启动子活力的改变会即时导致荧光素酶活性的改变，而荧光素酶不会积累。相反，CAT在哺乳细胞中的半衰期为50小时，在转录减弱后，酶的全部活性还能维持比较长的时间。④荧光素酶浓度在10^{-16}mol/L(10pg/L)到10^{-8}mol/L(1mg/L)范围内，荧光信号强度与酶浓度成正比。在理想条件下，可检测到10^{-20}mol/L的

荧光素酶,因此小量转染细胞裂解物就可以测定酶活性。

为了排除转染及蛋白质提取等过程中的误差对实验结果的干扰,目前,常采用双荧光素酶报告基因测试系统。双荧光素酶报告基因测试系统(DLR)灵敏、方便,在一个系统中用于测量两个单独的荧光素酶报告基因:萤火虫荧光素酶和海肾荧光素酶。将萤火虫荧光素酶作为实验报告基因,海肾荧光素酶作为对照报告基因。萤火虫和海肾荧光素酶都具有生物发光报告基因的卓越的测试特点,但它们在进化上的起源不同,因此,具有不同的酶学结构与底物要求。1986 年,萤火虫荧光素酶基因被用作测定基因表达的报告基因,在此之后获得了广泛的应用。萤火虫荧光素酶是一个 61 000 单亚基蛋白质,酶活力不需要翻译后修饰,在翻译后即可作为遗传报告基因。在 ATP、Mg^{2+} 和 O_2 的存在下,通过萤火虫荧光素的氧化反应而发光。海肾荧光素酶,一个 36 000 单亚基蛋白质,纯化自天然的 *Renilla reniformis*,含有 3%的糖类,但是和萤火虫荧光素酶一样,酶活力不需要翻译后修饰,在翻译后即可作为遗传报告基因。在海肾荧光素酶的催化下,海肾荧光素被氧化并释放光子,可用荧光检测仪对荧光作定量测定。

(五) DNA 与蛋白结合的研究

1. 凝胶滞留法(gel retardation assasy)　又称迁移改变法(mobility shift assasy),是检测 DNA 结合蛋白的一种简单迅速而又灵敏可靠的实验方法。凝胶滞留法的基本原理是:当 DNA 结合蛋白与相应的 DNA 片段或寡核苷酸结合而形成 DNA-蛋白质复合物后,使 DNA 片段的相对分子质量及电荷发生改变,因而在聚丙烯酰胺凝胶电泳中迁移率发生改变,通常较游离的 DNA 片段的泳动速率慢得多,在放射自显影 X 线胶片上形成一较游离 DNA 片段滞后的带型。此方法与其他实验方法相比,不但简单迅速,而且灵敏度高,可以检测出 fmol (10^{-15}摩尔)量的 DNA 结合蛋白。不同种类的 DNA 结合蛋白导致不同的迁移率,因此可以将与同一 DNA 片段结合的不同蛋白质分离开来。采用此方法可以鉴定特定基因调控序列中是否存在特定 DNA 结合蛋白的结合位点,也可以用来鉴定特定细胞核蛋白中是否存在某种基因的 DNA 结合蛋白,常用作为基因转录调控因子的初步筛选,是研究序列特异性 DNA 结合蛋白的最常用方法。

2. Southwestern 印迹　Southwestern 印迹是结合了 Western 印迹及 Southern 印迹两种试验方法特点而设计的一种检测序列特异性 DNA 结合蛋白的试验方法。其基本原理是,细胞核蛋白提取物经 SDS-聚丙烯酰胺凝胶电泳后,采用 Western 印迹方法将核蛋白条带转移到硝酸纤维膜上,然后与^{32}P-末端标记的 DNA 片段杂交。如果此核蛋白提取物中含有此种 DNA 片段特异结合的序列特异性 DNA 结合蛋白,则会在相应的蛋白质条带区域形成 DNA-蛋白质复合物,并可被放射自显影检测到。此方法的优点是可以测定到 DNA 结合蛋白的相对分子质量,因而为随后的蛋白质提取与纯化创造了条件。

3. 足纹法　足纹法是近年来发展起来的一种测定 DNA 结合蛋白在 DNA 上的精确结合位点的实验方法。事实上,足纹法的起源可以追溯到 20 世纪 60 年代末到 70 年代初。当人们用小球菌核酸酶轻微作用染色质后得到了长度均为约 200bp 或其倍数的 DNA 片段,这是由于 DNA 受到非特异性 DNA 结合蛋白-组蛋白的保护的结果。据此提出了核小体模型。这是 DNA 酶解足纹法的最早雏形。随后这一方法也应用到了序列特异性 DNA 结合

蛋白的结合位点鉴定上，Gilbert 等发现 *lac* 抑制子的结合可保护 *lac* 启动子免受甲基化试剂(DMS)的攻击，这是化学足纹法的雏形。早期的足纹法大多是将蛋白-DNA 复合物完全酶解，回收受到保护的 DNA 小片段，进而测定其序列，这种方法耗时又费力，直到 1978 年，Galas 等才以全新的方法发展了足纹技术，并正式提出了“足纹法”概念，他们发明的 DNase Ⅰ足纹法直至现在仍然是最常用的足纹法。

（张利宁　王　群）

第十五章　人类基因诊断与治疗

Chapter 15. Human Gene Diagnosis and Therapy

第一节　基因诊断

所谓基因诊断，原指基因水平上诊断遗传性疾病。这种技术后来也被用于诊断其他疾病。因此，基因诊断是从分子水平上对疾病进行的诊断。

一、基因诊断的方法

基因诊断的方法主要有 PCR 法、DNA 探针(probe)法和基因序列测定等，这些方法经常结合起来使用以方便对疾病的诊断。

(一) **PCR**(polymerase chain reacion)法

PCR 法就是多聚酶链式反应法。首先针对所要检查基因的 DNA 序列，设计并合成出两端的引物(primer)，再遵循 PCR 实验的步骤进行实验操作，从而扩增分离出特定的 DNA 片段，继之对扩增 DNA 片段进行电泳，从电泳所获得的不同带型，可以了解到该片段的异常情况，进行疾病诊断。由于 PCR 方法的成本低、快速和对样品质量和数量要求不高等特点，已开始用于多种遗传病的产前诊断(地中海贫血和血友病等)、传染病(如艾滋病)等。例如，血友病 A(HA)是一种常见的 X 连锁隐性遗传性出血性疾病，女性为携带者，男性发病，是由于 FⅧ基因 22 密码子倒位造成的，用长距离 PCR 的方法可检测出 FⅧ基因倒位的胎儿，以便进行产前诊断。首先设计合成三对特异性引物，PCR 扩增后，进行琼脂糖凝胶电泳，根据电泳条带的大小和相对分子质量的大小综合判定患者是否存在 FⅧ基因内含子 22 倒位。若只出现 12kb 带，判为野生型(非倒位)；若只出现 11kb 带，判为倒位患者；若出现 12kb 和 11kb 两条带，判为倒位的携带者。

(二) **RFLP**(restrictive enzyme fragment long polymorphogenesis)分析法

RELP 分析法即限制性酶片段长度多态性分析法。在非缺陷型疾病中，突变常常会产生新的或消除原有酶切位点，因此，可根据 DNA 限制性片段长度多态性不同进行基因变异分析，对疾病做出诊断。首先设计适当的引物，使扩增片段包含某一个或数个多态性的限制性内切酶位点，用 PCR 的方法对目的基因进行扩增后，用限制性内切酶对 PCR 产物进行消化，电泳后，根据正常人与病人这一等位基因酶切片段长度上的不同，再结合基因连锁分析，

可以了解某一特定基因上的变化。这一方法已在部分突变所致β-地中海贫血、血友病等疾病的诊断得到广泛的应用。

(三) DNA 探针法

DNA 探针是指用核素、酶、荧光分子或化学催化剂等标记特定 DNA。利用制备的特异 DNA 探针与标本进行杂交,通过杂交结果可确定检测标本中有无同源互补序列,从而检出所要查明的 DNA 或基因上的变化。

(四) 基因序列分析法

不管是 RFLP 法、DNA 探针法,还是 PCR 法,如要彻底了解相关基因的变化情况,均要对之进行 DNA 序列分析。其结果将清楚而明确地表明基因上的哪一个或哪几个碱基发生了突变。

二、基因诊断的疾病

当前,运用上述技术方法对遗传性疾病已能明确诊断(包括产前诊断)的有镰刀性贫血症、β-地中海贫血、Duchenne 肌营养不良症等 110 多种。能明确诊断的病毒、细菌、真菌及寄生虫引起的感染性疾病有腹泻病原菌、性病病原体、巨细胞病毒、乙型肝炎病毒、疱疹病毒、艾滋病毒(HIV)等约 40 种。应用抑癌基因 *p*16、*p*53 及基因重排分析等,可对某些白血病及恶性肿瘤进行基因诊断或产前诊断。对某些动脉粥样硬化易感性、职业病、中枢神经系统的疾病,也可进行基因诊断。

第二节 基因治疗

所谓基因治疗(gene therapy)是指将正常的外源基因导入生物体或人体细胞内,以弥补所缺失的基因或关闭、降低异常表达的基因,达到治疗疾病的目的。简言之,用基因治病的方法就叫做基因治疗。

世界上第一次真正在临床上进行基因治疗的是美国加州大学的一位科学家。他试图用基因治病的想法在当时几乎成了天方夜谭。他给两名晚期β-地中海贫血的病人导入了相关的基因,进行基因治疗,但最终失败。他因事先未履行必要的审批手续而给自己带来了很大的麻烦,但许多人称赞他是一位勇敢的开拓者。世界上第二次真正在临床上进行基因治疗的是美国国立卫生研究院(NIH)的 Blaese 和 Anderson。他们在 1990 年 9 月用重组反转录病毒表达腺苷酸脱氨酶(*ADA*)基因,将之导入一名 *ADA* 基因缺陷症女孩的 T 淋巴细胞,之后分批回注该重组淋巴细胞到患者血液中。3 个月后,回注体内的携带重组基因的淋巴细胞分泌了 ADA 酶,使这名患者先天缺损的免疫系统趋于正常。这是世界上基因治疗成功的第一个事例。之后又以基本相同的办法治疗了几位病人。这一成功事例激励了全世界的医学科学家,自此全世界风起云涌,掀起了基因研究治疗的狂澜。人类基因治疗最早着眼于遗传病,现在基因治疗除用于多种遗传病的治疗外,也用于治疗恶性肿瘤和其他疾病。随

着基因治疗基础研究的不断突破，今后临床基因治疗的疾病范围可望进一步扩大。

一、基因治疗的类型

根据受体细胞的不同，当代的基因治疗研究可分为性细胞(germ cells)基因治疗和体细胞(somatic cells)基因治疗。

(一) 性细胞基因治疗

顾名思义，性细胞基因治疗是通过基因治疗使非正常(缺损失活或过量表达某个或某些基因)的性细胞恢复正常，不仅治愈被治疗的病人，而且可纠正其错误的遗传背景，使其子孙后代不再罹患这一疾病。这是一个根治遗传性疾病的策略。主要是应用同源重组或基因打靶的方法，以实现目的基因对细胞染色体的定位整合或定点修复或定点突变，克服了盲目性，且可以遗传给后代子孙。但其效率很低，技术难度较大，以至当前还未能进入临床实验。

(二) 体细胞基因治疗

体细胞基因治疗是针对具体患者所存在的基因缺损、失活或基因过量表达的实际情况，把相应的正常基因导入患者的体细胞，以治愈被治疗的具体患者，但不能纠正其错误的遗传背景。故这个疗法被认为是一个“治表而不治本”的疗法。如，导入 *ADA* 基因可以使 *ADA* 基因缺陷的患者获得足够的 *ADA* 酶而痊愈；但这一患者所缺损的 *ADA* 基因并未得到修复，这样，其子孙后代中仍会有人罹患 *ADA* 基因缺陷症。然而，体细胞基因治疗的优越性是不可否认的。首先，他使一向认为不可治愈的遗传性疾病变成可以治愈的疾病，这是一个历史性的突破。其次，即使是每 3 个月治疗一次，也比天天打针吃药的传统办法简便，经济得多，更主要的是方便了病人，应当认为这是临床医学上的一大进步。

二、基因转移系统

目前已有多种基因转移方法可将目的基因导入受体细胞，对基因治疗较有意义的载体有反转录病毒(RV)、腺病毒(Ad)、腺病毒伴随病毒(AAV)、单纯疱疹病毒(HSV)、脂质体与受体介导的蛋白(RMP)。在迄今进行的临床基因治疗中，大部分采用反转录病毒载体介导的基因转移。

(一) 反转录病毒

反转录病毒被选作体细胞基因治疗的载体，是因为反转录病毒可进入细胞并整合到染色体上，病毒基因在宿主细胞内可适宜表达但不影响病毒生活周期，但从受体细胞出芽的病毒粒子给基因治疗带来安全性问题。不具包装功能的病毒载体与包装细胞系的应用，使病毒出芽的可能性极大地降低。

RV 前病毒 DNA 可分成两部分，即顺式作用与反式作用序列(图 15-1)。顺式作用序列包括两端长末端重复序列(long terminal repeat sequence，LTR)、包装信号 Ψ 位点等，为病毒

复制的必需调控区；反式作用序列编码包装蛋白 gag、pol 与 env。构建载体时去除反式作用序列，插入外源基因，再与质粒重组以利扩增。重组的病毒基因组是复制缺陷型的，只能产生病毒 RNA，不编码病毒的结构蛋白。

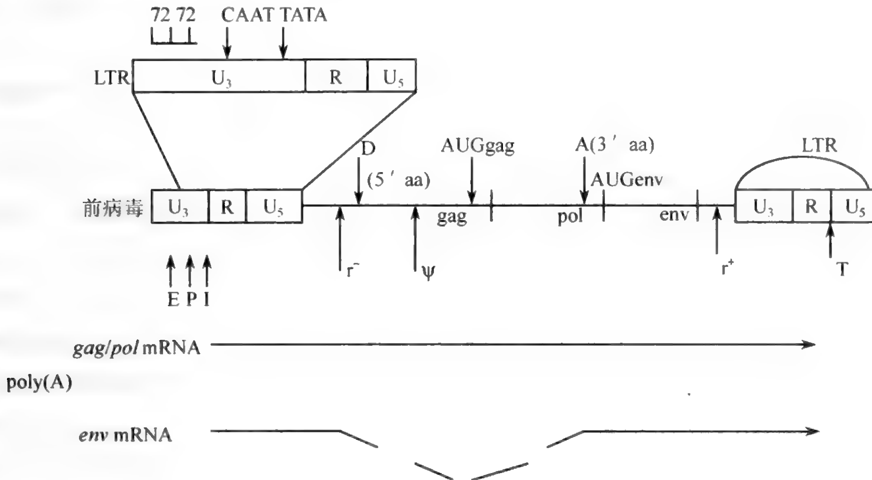


图 15-1　鼠莫洛尼白血病反转录病毒前病毒 DNA 结构

E，增强子；P，启动子；I，病毒 RNA 合成起始部位；D，给位裂解部位；r⁻，副链 DNA 复制起始部位；Ψ，包装信号；A，主要受体裂解部位；r+，正链 DNA 自制起始部位；T，病毒 RNA 合成中 3′多聚腺苷酸加成部位；LTR，长末端重复；U_3、U_5、R，长末端重复的部分；gag，病毒核心特异抗原；pol，依赖 RNA 的 DNA 聚合酶（反转录酶）；env，外壳蛋白。72.72，72kb 重复；CAAT、TATA，启动子区保守盒

1．载体的种类　重组的反转录病毒载体大致可分为四类：第一类是双表达型载体，含两种外源基因，分别取代 *gag/pol* 片段与 *env* 基因，双表达载体的特点是病毒基因的调控区与外源基因的表达密切相关；第二类是内部含启动子的病毒载体，筛选基因紧接长末端重复序列后面，其表达受 5′LTR 启动子控制，目的基因前附有一个外加启动子，其选用由目的基因与靶细胞而定，与双表达载体相比有较大的自由度，不足之处是内部启动子的存在可能会影响病毒的效价；第三类是自身灭活载体，3′LTR 缺失一段序列，由于 3′LTR 是两端 LTR 的模板，病毒复制时其缺失，这样病毒顺式结构不会干扰外源基因的表达，减少了 LTR 激活相邻癌基因的可能性。其缺点是病毒滴度特别低。第四类是双拷贝载体，其主要特点是外源基因插入 3′LTR 的 U_3 区，在转染细胞内基因被复制并转到 5′LTR。外源基因的新位置处于反转录单位之外，减少了该转录单位对外源基因表达的负影响，使插入基因的调控序列较正常地发挥作用。体细胞基因治疗多采用内部含启动子的病毒载体。

2．包装细胞　上述 4 种缺失编码包装蛋白序列的病毒载体，必须提供包装蛋白以完成病毒包装。将缺失包装信号 Ψ 序列的反转录病毒 DNA 导入细胞系即构建成包装细胞系，这种细胞产生包装蛋白，因自身缺失包装信号不能包装。病毒载体导入细胞后，含有包装信号的载体与包装细胞作用互补共同完成病毒包装。出芽的病毒粒子转染目的细胞后，因目的细胞不含编码包装蛋白的序列，病毒 DNA 不能包装出芽，从而避免了扩散。

（1）第一代包装细胞系：Ψ2 是单向性包装细胞系，此类出芽的病毒粒子只能感染啮齿

类动物,这种属特异性是由病毒粒子表面的糖蛋白类型所决定。这种包装细胞系的缺点是易产生辅助病毒,因为反转录与整合的信号是完整的,当导入的反转录病毒载体与缺失的3′末端有交联时,可产生辅助病毒,受感染细胞的内源序列亦可修复病毒的基因组缺陷。如辅助病毒的基因组同时被包装,病毒粒子进入受体细胞后可进一步感染其他细胞,即包装信号的去除仅能降低但不能排除辅助病毒的产生。

(2) 第二代包装细胞系:以 PA317 为代表,PA317 细胞株是双向性包装细胞系,由 PA317 出芽的病毒粒子,不仅可感染啮齿类细胞,也可感染包括人类在内的灵长类细胞。它不仅缺失包装信号 Ψ 序列,还缺失 5′LTR 的 5′末端,同时 3′LTR 与负链的起点位点被 SV40 的聚腺苷位点取代。经过如此处理,即使这一结构的 RNA 也被包装,仍不能反转录。比起单纯缺失 Ψ 序列的包装细胞系,PA317 产生辅助病毒的几率要低得多。

(3) 第三代包装细胞系:将 *gag/pol* DNA 片段与 *env* DNA 片段分离并引入突变与缺失,于 *gag* 或 *env* 两个区域的转录位点分别插入连接子(linker)以产生突变,大部分分离的转录单位的 LTR 被 SV40 的聚腺苷信号取代,由此构建单向性包装细胞系 ΨCRE 与双向性包装细胞系 ΨCRIP,这两个细胞株可看成第三代包装细胞系。在包装细胞的两个转录单位与载体或内源成分共 3 个 DNA 重组才能产生辅助病毒,但因突变位于转录单位,可能仅需两种重组事件发生即可产生辅助病毒。重组发生于包装细胞内的病毒序列与反转录病毒载体或内源成分的反转录样结构。实验证明当采用特定的载体时,ΨCRE 与 ΨCRIP 细胞系产生辅助病毒的倾向小于 PA317 细胞。

(二) 腺病毒

腺病毒(Ad)的病毒粒子无包膜,二十面体对称,基因组为线状双链 DNA。腺病毒易于培养和纯化,基因组的复制与转录依赖于细胞的聚合酶,因此病毒基因的表达调控可以模拟宿主细胞的基因表达机制,允许插入载体的外源基因长度可达 36kb,病毒滴度较高(10^{11} pfu/ml)。能感染非分化细胞,可原位感染组织,尤其是肺。其主要缺点是目前所用的载体含较多的腺病毒序列,可能刺激免疫反应或有其他副作用。此外,因病毒载体未整合在染色体 DNA 上,基因表达可能不稳定。

(三) 腺病毒伴随病毒

腺病毒伴随病毒(AAV)亦称依赖性病毒(dependovirus),它的繁殖依赖于辅助病毒,后者包括腺病毒和疱疹病毒。AAV 成熟病毒粒子内,或含有正链 DNA,来自不同病毒粒子的正负链 DNA 在试管内可形成双链 DNA 结构。AAV 与 Ad 的混合感染可以明显地抑制 Ad 的复制。在这种溶细胞性感染中,复制需要 Ad 或 HSV 的共同感染。当缺乏辅助病毒时,AAV 感染导致其基因组整合到宿主细胞染色体基因组上,形成潜伏感染状态,然后当有辅助病毒感染同一细胞时,AAV 基因组 DNA 将复制。取自潜伏感染细胞的染色体与 AAV DNA 的原位杂交证实,AAV 基因组整合的位点在 19 点染色体 q13.4。数个独立的整合位点的顺序分析显示,断裂点均位于宿主细胞 DNA 上 100bp 内,将基因组 DNA 特异地整合到宿主细胞染色体上单一位点以形成病毒潜伏感染。在迄今所见的病毒 DNA 整合中,AAV 的这种位点特异性整合是仅有的。

（四）单纯疱疹病毒

HSV 病毒是有包膜的双链 DNA 病毒，线形 DNA 链中有末端重复序列和内部重复序列，链中有缺口。HSV 基因组中有许多基因在细胞培养中对病毒生长是非必需的，所以，HSV 载体可导入大片段基因以至多个基因，构建 HSV 载体期望得到有组织特异性的、稳定、适宜的表达。

（五）脂质体

用一定比例的磷脂酰胆碱与磷脂酰丝氨酸共溶于氯仿，制备有双层磷脂膜的脂质体。将待转移的 DNA 分子包入脂质体中，受体细胞可吞噬这种含外源基因的脂质体，以此达到基因转移的目的。脂质体已用于直接体内基因转移。将对靶细胞特异的单克隆抗体包埋于脂质体的脂质双层，用超声处理将目的基因 DNA 包入这种"免疫脂质体"中。DNA、免疫脂质体的腹腔注射使基因特异性转移至胞膜对单克隆抗体产生免疫反应的细胞。

（六）受体介导的蛋白

受体介导的蛋白质研究较多的是唾液酸糖蛋白与转铁蛋白。转铁蛋白受体介导的基因转移已成功地将 DNA 转移至多种离体细胞，唾液酸糖蛋白受体介导的基因转移对肝细胞高度特异，体外与体内实验均已将 DNA 转移至肝细胞。蛋白质与多聚赖氨酸或其他阳离子化合物偶联，所带正电荷有利于形成 DNA/蛋白复合物，复合物经静脉注射后，受体所在的部位将决定其细胞特异性，影响核内体/溶酶体旁路的药物（如氯喹）增加基因转移的效率。蛋白受体介导的基因转移属直接体内的方法，这一方法已显示出对肝细胞的突出优点，不仅是唾液酸糖蛋白介导的基因转移对肝细胞高度特异，而且不像反转录病毒载体那样需要做肝部分切除才能实现高效转移。受体介导的蛋白质携带反义 DNA，可显著抑制肝炎病毒基因的复制与表达。

三、受体细胞

不管是直接体内还是间接体内基因转移，选择目的基因表达的组织细胞最好是组织特异性细胞，在一种组织细胞中表达后可分泌或以其他方式进入靶细胞的也可应用。对于间接法基因转移，还应考虑：①细胞较易从体内取出且有增殖优势，最好生命周期较长；②离体细胞较易受外源遗传物质转化；③细胞经转化和一定时间培养后再输回体内仍能成活。

（一）造血细胞

造血细胞种类多样且在体内多处分布，有多种疾病累及造血细胞，造血细胞较符合间接体内基因治疗应考虑的因素，因此造血细胞作为基因转移的受体细胞研究较多，其中骨髓最受重视。因骨髓含多种造血细胞且处于不同发育阶段，如基因转移至分化细胞，则转移基因只能在这种特定细胞的生命周期内表达。只有将基因转移至骨髓多能干细胞，才能使转移基因终生存在于所有造血细胞内。目前主要采用反转录病毒载体介导的基因转移，通过包

装细胞包装出重组病毒,在转染造血细胞后,进行短期培养,再输回动物体内。这种方法类似自体组织移植,其主要区别是移植细胞受到反转录病毒的转染。主要问题是转移系统的安全性与有效性,转移基因在受体细胞表达的适宜性。

在转染目的细胞之前,对导入外源基因的双向性包装细胞系克隆作系统筛选,是为了最大限度地克服由反转录病毒载体的不稳定性所造成的基因缺失、重排和病毒滴度低等缺点。检测受转染动物外源基因的整合与表达,并用非常敏感的方法检测有否辅助病毒的产生,并从整体水平上考虑安全性与有效性。反转录病毒载体仅介导外源基因转移处于分裂状态的细胞。因此这种方法适合于分裂较快的血细胞,也可用于皮肤成纤维细胞、肝细胞等;不适合于增殖较慢甚至不分裂的细胞,尤其是肌肉细胞与脑细胞。已有许多实验室可分离鼠骨髓干细胞及人类与其他高等动物的骨髓干细胞。如转移后 4 个月仍能在受体动物的骨髓和(或)淋巴组织内检测到转移基因的表达,即认为转移基因已整合于造血干细胞。应用这种方法,已在鼠体内得到腺苷脱氨酶基因的长期表达,二氢叶酸还原酶突变株基因已在鼠体内表达并显示对转染细胞的选择优势。反转录病毒 DNA 在转移细胞染色体基因组上随机高效整合,但重建的动物造血组织细胞的基因及 DNA 中反转录病毒的不同整合位点数较少,说明只有少数受感染细胞在骨髓细胞内成活。转染及筛选过程使骨髓增殖能力降低,可能是丢失干细胞或对干细胞转染不足的原因。对人类基因治疗需要高效基因转移。骨髓细胞培养、造血干细胞的纯化,在细胞培养中使用多种生血因子联合应用,可望增加基因转移效率。

(二) 肌肉细胞

反转录病毒载体介导的基因转移对骨骼肌细胞几乎无效。骨骼肌细胞的突出特点是可采用直接体内基因转移,即将 DNA 直接注射入骨骼肌细胞。用含 β-半乳糖苷酶基因的质粒直接注射入小鼠骨骼肌细胞,在注射区域的肌细胞检测到该基因的表达。用同样的方法注射含荧光酶基因的质粒,在 2 个月的检测中荧光酶基因以较稳定的水平表达,提示其表达持续时间会更长。用金粒子携带 β-半乳糖苷酶基因轰击鼠体内肌肉细胞,在轰击后 2 个月内鼠肌肉内有该基因的存在与表达。直接注射与轰击后外源 DNA 似乎以未整合的染色体状态存在,直接注射法仅对骨骼肌细胞有效,而轰击适用于多种细胞。

血管平滑肌细胞可以培养,经遗传操作后再送回血管。反转录病毒载体携带人 *ADA* 基因与细菌 *NEO* 基因转染培养的兔平滑肌细胞,用特制的球形导管剥去内皮细胞后,将受转染的兔平滑肌细胞送回动脉壁。人 *ADA* 基因在兔血管壁上以和内源 *ADA* 相近的水平持续表达半年以上。

(三) 成纤维细胞

用携带目的基因的反转录病毒载体转染成纤维细胞,已表达具有生物活性的 ADA、葡萄糖脑苷酯酶、嘌呤核苷磷酸化酶、低密度脂蛋白受体与凝血因子Ⅸ等。经过遗传操作的成纤维细胞的胶原基质可分泌因子Ⅸ,将这种基质移植到动物皮下检测到因子Ⅸ的表达。这一方法已用于临床治疗 B 型血友病。

经遗传操作的成纤维细胞株,能在受体细胞形成肿瘤。生长激素表达随时间推移而增加,说

明移植细胞持续增殖。正常成纤维细胞行为完全不同，移植后能阻遏转移基因的表达，值得探讨的是，正常成纤维细胞能否在移植后持续表达一定水平的外源基因及对大动物是否有效。

（四）肝细胞

肝脏是重要的代谢器官，有多种遗传病由肝脏特异的基因产物缺陷引起，且肝脏是肿瘤的高发部位，因此肝细胞成为基因治疗研究的重要目的细胞。外源基因转移至体内肝脏的方法有两种，直接体内基因转移法和间接基因体内转移法。间接基因体内转移是由反转录病毒载体介导将外源基因转移至离体细胞，然后做自体组织移植。动物实验证明，转基因鼠的肝细胞经同种异体移植可以成活并继续发挥肝细胞的功能。直接体内基因转移是通过受体介导使重组病毒载体直接导入肝细胞，从而将外源基因转移至体内肝脏。将唾液酸糖蛋白与多聚赖氨酸偶联通过与肝的唾液酸糖蛋白受体相结合转移大肠杆菌β-半乳糖苷酶基因至鼠肝细胞，约有0.1%肝细胞被转染并检测到低水平的β-半乳糖苷酶。这种受体介导的基因转移系统与腺病毒的核内溶胞能力相结合，可促进DNA转移至肝细胞并产生高水平基因表达。

（五）其他细胞

用反转录病毒载体介导的基因转移至离体的猪与犬内皮细胞，自体内皮细胞移植后，外源基因表达可持续1～2个月。内皮细胞分泌抗凝血因子以抑制所在部位血栓形成，血管内皮细胞只有一层，广泛分布、数目有限及局部作用的特点限制了这一实验的实用性。

肿瘤浸润的淋巴细胞（TIL）以专门攻击肿瘤为其突出的特点。在体外用反转录病毒载体转染TIL，扩增后再输回患者体内，受转染的TIL在患者体内尤其是肿瘤部位持续存在，未见明显副作用。细胞携带肿瘤坏死因子基因，分泌的肿瘤坏死因子可杀伤肿瘤细胞。

四、基因治疗的应用

在*ADA*基因缺陷病基因治疗取得成功的激励下，世界各国迅速开展了广泛的研究。下面是当时开展研究的一些情况（表15-1、15-2）。

表15-1　美国基因治疗的疾病（29种）

产品名	公　司	适应证	开发阶段
AAVCFTR	Targeted	囊性纤维变形、窦炎	Ⅰ期临床
AD-P53	Introgen Therapeutics	非小细胞型肺癌 头颈部癌症	Ⅰ期临床
抗HIVT细胞治疗	Cell Genesys	HIV感染	Ⅰ/Ⅱ期临床
CFTR腺病毒载体	Genetics Therapy	囊性纤维变性	Ⅰ/Ⅱ期临床
CFTR基因治疗	Genzyme	囊性纤维变性	Ⅰ期临床
抗化疗骨髓	Genetize	化疗癌症患者	Ⅰ期临床
胞嘧啶脱氨酶基因-酶病毒载体	GenVce	结肠癌	Ⅰ期临床
DA/Hu(r).15转导自体肿瘤细胞;ITAT	Chiron Viagene	转移性肾细胞癌	Ⅰ期临床

续表

产品名	公　司	适应证	开发阶段
DA/Hu(r).15 转导自体肿瘤细胞;ITAT	Chiron Viagene	扩散的恶性黑色素瘤	Ⅰ期临床
DA/Hu(r).15 转导自自体肿瘤细胞和表达 γ-干扰素的自体肿瘤细胞	Chiron Viagene	扩散的恶性黑色素瘤	Ⅰ期临床
DA/Hu(r).15 转导自体肿瘤细胞;ITAT	Chiron Viagene	成神经细胞瘤	Ⅰ期临床
hIFN-g 反转录病毒载体	Chiron Viagene	转移性黑色素瘤	Ⅰ期临床
葡糖脑苷酯酶基因—反转录酶病毒载体	Genetic Therapy	Gaucher 病	Ⅰ期临床
GV-10	GenVec	囊性纤维变性	Ⅰ期临床
HIV-IT I II B env/re 反转录病毒载体	Chiron Viagene	无症状 HIV 感染	Ⅰ期临床
HIV-IT(V)Retro Vector™转导的自体成维细胞	Chiron Viagene	无症状 HIV 感染	Ⅰ期临床
HIV-IT(TAF) HS-tk	Genetic Therapy	成人脑瘤	Ⅰ/Ⅱ/Ⅲ期临床
		儿童脑瘤	Ⅰ/Ⅱ/期临床
		寄生虫病	Ⅰ期临床
IC-imm 反转录病毒载体	Targeted Genetics	HIV 感染	Ⅰ期临床
IL-2 基因修饰的肿瘤	RPR Gencell	乳腺癌	Ⅰ期临床
Lg-GC 反转录病毒	Targeted Genetics	Gaucher 病	Ⅰ期临床
LN-CTL 反转录病毒		HIV 感染	Ⅰ期临床
MDR 反转录病毒载体	Genetech Therapy	乳腺癌	Ⅰ期临床
MDRxL™	Ingenex	保护造血干细胞免受化疗毒性	Ⅰ期临床
携带 TNF 反转录病毒载体	Chiron	黑色素瘤	Ⅰ期临床
RGG-0853, ELA (脂复合物)	Chiron	卵巢和乳腺癌	Ⅰ期临床
RV-AS-KRAS	Introgen Therapeutics	非小细胞型肺癌	Ⅰ期临床
RV-p53	Introgen Therapeutics RPR Gencell	非小细胞型肺癌	Ⅰ期临床
tgAAVCF	Targeted Genetics	囊性纤维变性	Ⅰ/Ⅱ期临床

表 15-2　美国反义治疗产品(6 种)

产品名称	公　司	适应证	开发阶段
GEM91	Hybridon	HIV 感染,AIDS	Ⅱ期临床
ISIS2302	Isis Pharmaceutical	炎症性疾病	Ⅱ期临床
ISIS2922	Isis Pharmaceutical	AIDS 患者的巨细胞病毒(CMV)视网膜炎	Ⅲ期临床

续表

产品名称	公　司	适应证	开发阶段
ISIS3521	Isis Pharmaceutical	癌症	Ⅰ期临床
ISIS5132	Isis Pharmaceutical	癌症	Ⅰ期临床
LR-3001	Lynx Therapeutics	慢性骨髓性白血病进展期	Ⅰ期临床

（一）遗传性疾病的基因治疗

这一方面的研究多半属于单基因缺陷所引起疾病的基因治疗，主要有①血友病：Ⅷ-C(促凝成分)和 因子缺陷症；②免疫缺陷病：腺苷酸脱氨酶(ADA)缺陷症、嘌呤核苷磷酸酶(PNA)缺陷症；③贫血：β-地中海贫血、镰刀型红细胞贫血；④肺气肿：α_1-抗胰蛋白酶缺陷症；⑤溶酶体储藏障碍：葡萄糖苷酯酶缺陷症(Ganoher 病)；⑥其他代谢病：如苯丙氨酸化酶(PAH)缺陷症、次黄嘌呤鸟嘌呤磷酸核糖转移酶(HGP)缺陷症等。

（二）恶性肿瘤的基因治疗

在美国和德国科学家进行了大量的预备性实验后，美国 NIH 的科学家进一步构建重组的 TIL(肿瘤浸润淋巴细胞)，它们能表达 100 倍于正常水平的 TNF。1990 年 11 月，他们获准用表达 TNF 和 TIL 治疗黑色素瘤，临床实验的人数由开始的 2 例增至 50 例。另外，用表达 IL-2、IFN-2a 和 IL-1 的 TIL 治疗神经细胞瘤及白血病等的工作也开展了起来。美国还应用反转录病毒将毒素基因(蓖麻毒素和骨髓灰质炎病毒毒素中所含的一种蛋白酶的基因)导入癌细胞内，在靶细胞内表达毒素并挥发杀伤作用，但对其他细胞毒性较低。而且这类载体已被改造成只攻击带有特殊标志的细胞，也可运送常规的治疗蛋白。该研究在基因治疗上发挥了分子导向的思想，给人以启示。

（三）传染病的基因治疗

对于不治之症艾滋病，美国科学家用重组反转录病毒将 CD4 分子导入人上皮细胞及小鼠成纤维细胞使之表达，试图抑制艾滋病病毒 HIV。类似的方法还被用于治疗人乳头瘤病毒、巨细胞病毒及肝炎病毒的感染。此外，还研究了一种药物作用系统，将一段酶序列基因插入受感染细胞，所表达的酶可催化无活性的药物原体转化为有毒性作用的形式，从而有选择地杀伤受感染细胞以避免毒性对全身的作用。还有将编码可使艾滋病毒 HIV RNA 降解的核酶基因转染入淋巴细胞，从而抑制 HIV 的传播。又将 HIV LTR-2′,5′寡腺苷酸合成酶杂合基因转染 HeLa-T4$^+$ 细胞，结果经修饰的带有 CD4 受体的这种细胞具有抗 HIV 感染的作用。

（四）其他基因治疗

在糖尿病治疗研究中，日本科学家将表达人胰岛素的组织培养细胞植入患糖尿病小鼠的腹腔内，结果植入的细胞存活并分泌胰岛素，使小鼠血糖水平明显下降。用基因治疗技术治疗肌营养不良、骨质疏松、类风湿关节炎及中枢神经系统疾病等研究亦在进行中。

五、基因治疗的现状和展望

多年来的探索,基因治疗取得较为满意的结果,除了 *ADA* 基因缺陷病之外,还有 *LDLR*(低密度脂蛋白受体基因)缺陷症、B 型血友病等。对于前者主要办法是导入 *LDLR* 正常基因并使之表达,弥补该受体的缺如,进而使 *LDL* 经受体结合与介导而被降解。对于后者,主要是导入病人所缺少的 *IX* 因子基因并使之表达,从而取得疗效。应用 *TK* 基因直接导入脑瘤病灶,也有较好的效果。

当前普遍认为,除某些单基因遗传病的基因疗法效果较好外,其他基因治疗的效果不佳。主要原因是对于有关疾病的发生机制了解尚属粗浅,治疗方案在技术上也待改进。

据认为,原来所阐明的遗传性疾病,大多数是单基因缺陷的疾病。原以为与遗传基因无关的疾病,当前认为,不仅与一个基因有关,而且往往同多个基因有关,甚至同某一个或某几个基因网络相关,并被称为多基因疾病,如恶性肿瘤、心脑血管疾病、自身免疫病、内分泌疾病等等。因此,在考虑基因治疗的方案时,事先需要考虑的问题,其涉及面将深广得多。如家族性高血压,当前已知与多达 30 多个基因有关。而且,有的与肾素-血管紧张素-醛固酮代谢系统相关;有的则与糖代谢系统相关,如胰岛素受体缺陷,有的则与循环系统,甚至与神经系统相关。在大约 6000 种遗传性疾病中,单基因疾病仅占 70%。基因治疗是一种基于分子遗传学的生物高技术,对于相关的基础理论知识如果掌握得不够,也就很难设计出可以达到十分满意疗效的技术方案。

对恶性肿瘤进行基因治疗的探索最为广泛,总的看来,疗效不能令人满意。具体的治疗方案可以分为两类。一是杀灭肿瘤细胞的基因疗法。采取这一疗法时,一般都进行基因的靶向性设计。但除了导入 *p53* 抗癌基因在某些情况下有一定的疗效之外,大部分的临床实验结果不理想。而且,可能用于靶向性设计的,能够强烈杀灭癌细胞的可用基因并不多。另一种办法是把可提高机体免疫力的基因导入体内,以通过提高机体的免疫力而控制肿瘤的发展或逐步缩小肿瘤。但往往在动物模型上进行的试验效果较好,同样的方案用于肿瘤病人的临床试验,则效果不佳。可能是因为动物实验的荷瘤动物,其免疫系统尚未遭到严重的破坏,以提高免疫力为目标的基因疗法能够得到较好的结果。而患有恶性肿瘤的病人,往往其免疫系统已经遭到了比较严重的破坏,以致采取提高免疫力的疗法,不能取到如期的效果。因此,有人主张,在采用这一疗法之前,要首先研究具体治疗对象本身免疫系统的实际状况,再来分析与决定采取提高免疫力的基因疗法是否可取。这就是说,提高免疫力的基因疗法要个体化。

目前,已应用的基因治疗方案,在理论和技术上还存在不少问题。为了推动基因疗法的发展,强化基础理论研究与改进已有技术的研究,都是十分重要的。为此,在大力推进人类基因组的研究计划中,研究疾病、特别是多基因疾病与相关基因的关系的命题悄然兴起,而且研究得如火如荼。在技术研究方面,已经发展出多种目的的基因载体,各有优缺点。而且,着力发展非病毒载体的基因疗法。

总的看来,基因治疗的探索尚处在初步阶段,通向真正成功的道路是漫长的。但它标志着临床医学的一次历史性变革。人们预期,它将可能在不远的将来,成为临床医学中的一种常规疗法,前景光明,道路曲折。

(张利宁　丁培芳)

第十六章　DNA 多态性

Chapter16. DNA Polymorphism

随着人类基因组计划(human genome project,HGP)的实施和完成,人类 DNA 序列已基本清楚,并由此延伸出许多新的研究分支,如功能基因组学(functional genomics)和基因组序列变异的研究。

基因组序列出现变异是基因组在遗传学上的基本特征之一,它是人类进化和适应环境的必然结果。基因组中的这种变异,赋予生物医学家以崭新的思想,从而从这些差异中发现了人类之中各种生物学现象的奥秘,如健康状况、对疾病的易感性、寿命的长短、人类的起源等等。这些对于人类认识自己的遗传背景和据此采取保护措施具有十分重要的意义。

研究基因组差异如果直接从大规模全序列比较显然难以做到,但可从均匀分布于基因组的具有代表性差异位点开始,密度不断增加,最终找到存在于所有人中的各种 DNA 序列差异,这就是 DNA 多态性(DNA polymorphism)。

DNA 多态性的研究始于 20 世纪 80 年代,大致可分为三个阶段:限制性片段长度多态性、微卫星 DNA 和单核苷酸多态性。

第一节　DNA 多态性研究概述

一、限制性片段长度多态性

RFLPs (restriction fragment length polymorphisms) ,即限制性片段长度多态性,是最早用来研究 DNA 多态性的方法。利用 DNA 限制性片段长度多态性的基因连锁分析,主要是应用 DNA 内切酶对不同 DNA 样品中的等位基因进行酶切分析,不同个体之间等位基因被酶切的结果在长度上不同,再结合基因连锁分析,可以了解某一特定基因在人群中的变化。1987 年,第一张较完整的人基因连锁图绘制成功,共确定了 500 多个 RFLP 位点,并定位了 Huntington 病、CF(囊性纤维化)和多个癌相关基因。

二、微卫星 DNA

微卫星 DNA 是广泛存在于原核及真核生物基因组中的一种串联重复序列,其中,近年发现的短串联重复序列(short tendem repeats),和可变数目串联重复序列(variable number tandem repeats ,VNTRs),已发展为一种新型的 DNA 高度多态性遗传标记系统。

STRs 和 VNTRs 的共同特征是:由 2～70bp 的核心序列(基本单位)呈串联重复多次,

中间没有间隔，总长度一般小于 1kb。但二者的核心序列长度不同。

STR，又称为微卫星 DNA（mini-satellites DNA），其基本单位是 1bp～8bp 的串联重复，每个微卫星 DNA 的核心序列结构相同，重复单位数目约 15～60 个。目前已知的 STR 有 8000 多个，主要形式为：(CA)*n*，(GA)*n*，(AA)*n*，(GG)*n*，(CAA)*n*，(CGG)*n*，其中以(CA)*n*，(GT)*n* 为多见。STRs 广泛存在于人类基因组，约占 5%，其多态性主要来源于串联数目的不同。1992 年，Welssenbanch 等以 STR 为标记制作第二代连锁图取代了 RFLP 连锁图。

VNTR，又称为小卫星 DNA（small-satellites DNA）。每个 VNTR 都含有 10～15bp 核心序列，这类重复可能来源于减数分裂期间的同源染色体或染色体内部的不对等交换。VNTR 与 STR 比较类似，但是重复顺序更长；多态性及分布方面不如 STR，因而 VNTR 作为遗传标记只是一种过渡，目前已较少使用，但是，VNTR 曾经对 STR 的基因定位运用起到推动作用。

作为一种新型的 DNA 多态性遗传标记，微卫星 DNA 具有种类多、分布广泛、高度多态性、杂合性高、重组频率低的特点，有高度的个体特异性，遵循孟德尔共显性遗传定律，随着它的广泛发现和应用，使得人们加快了疾病相关基因定位、分离、克隆和筛选的研究步伐。

三、单核苷酸多态性

单核苷酸多态性（single nucleotide polymorphism，SNP）是人体 30 亿个碱基对在排列顺序中出现的个体差异。在不同人的同一条染色体和同一个位置的核苷酸序列中，可能绝大多数核苷酸排列序列一致，只有一个地方的核苷酸序列不同，这种现象就是“单核苷酸多态性”。

在 DNA 多态性的不同表现形式（如点突变、插入、缺失和不同数目串联重复等）中，以单核苷酸多态性发生频率最高，它是人类可遗传的变异中最常见的一种，占所有已知多态性的 90% 以上。1996 年，Lander 成功地用单核苷酸多态性作为遗传标记制备第三代遗传连锁图。

随着对 SNP 检测和分析技术的进一步发展，尤其是与 DNA 芯片等技术的结合，它已成为第三代遗传标记，SNP 的深入研究可以为新的诊断方法、治疗方法提供基础，同时也是群体研究中探索及说明新基因功能的一个极其有效的工具。SNP 可以满足对疾病相关基因定位研究的需要，尤其是对多基因遗传病高精度基因定位的要求，并将最终取代目前最常用的微卫星标记技术进入基因应用研究的领域，对药物基因组学、诊断学和生物医学领域的研究起重要作用。

第二节 单核苷酸多态性

一、单核苷酸多态性的概念

单核苷酸多态性，指基因组序列中的由单个核苷酸的变异所引起的 DNA 序列多态性。

例如一个 SNP 可能把 DNA 的序列 AAGGCTAA 改变为 ATGGCTAA。人类基因组的 30 亿个碱基中，许多 SNP 不影响细胞功能，但科学家认为有些可能会使人易患病或影响对药物的敏感性。

SNP 所表现的多态性只涉及单个碱基的变异，一般只有两种碱基组成，所以 SNP 是一种二态的标记，即二等位基因（biallelie），又称为 SNP 的二态性。SNP 的这种变异多由单个碱基的转换（transition）或颠换（transversion）引起，其中转换与颠换之比大约 2∶1。SNP 在 CG 序列上出现最为频繁，而且多是 C→T，原因是 CG 中的 C 即胞嘧啶常甲基化，甲基化的胞嘧啶自发脱氨后即成为胸腺嘧啶。

在基因组 DNA 中，任何碱基均有可能发生变异，因此 SNP 既有可能在基因序列内，也有可能在基因以外的非编码序列上。位于编码区内的 SNP，称为编码区 SNP（coding SNP，cSNP）。尽管 cSNP 比较少，但它在遗传性疾病研究中却具有重要意义。

cSNP 分为同义 cSNP（synonymous cSNP）和非同义 cSNP（non-synonymous cSNP）两种。同义 cSNP 中发生无意义突变，即突变碱基与未突变碱基的含义相同，因而同义 cSNP 突变导致的编码序列的改变并不影响其所翻译的蛋白质的氨基酸序列。非同义 cSNP 中发生了有意义碱基突变，使得其编码蛋白质的氨基酸序列发生改变，从而影响了蛋白质的功能。这种改变常是导致生物性状改变的直接原因。cSNP 中约有一半为非同义 cSNP。

二、SNP 作为遗传标记的优势

SNP 自身的特性决定了它比 RFLP 和微卫星多态性标记更适合于对复杂性状与疾病的遗传解剖以及基于群体的基因识别等方面的研究，SNP 的应用对群体遗传学、制药业、法医学、癌症及遗传性疾病甚至进化的研究都将产生不可估量的影响。

（1）SNP 数量多，分布广泛。据估计，人类基因组中每 1000 个核苷酸就有一个 SNP，人类 30 亿碱基中共有 300 万以上的 SNP。

（2）SNP 易于进行自动化分析，适于快速、规模化筛查，缩短了研究时间。由于 SNP 具有二态性，非此即彼，在基因组筛选中 SNP 往往只需有或无的分析，而不用分析片段的长度，这就利于发展自动化技术筛选或检测 SNP。

（3）SNP 具有高度的稳定性，研究表明，单核苷酸多态性存在于生殖细胞中，并可以全部从上一代传给子代，因而不会因突变而造成对人群遗传分析的困难。

（4）部分位于基因内部的 SNP 可能会直接影响产物蛋白质的结构或基因表达水平，因此，它们本身可能就是疾病遗传机制的候选改变位点。

三、单核苷酸多态性的研究意义

单核苷酸多态性是人类基因组 DNA 序列变异的主要形式，在人类基因组中广泛存在，具有很高的信息含量，与人类性格、身高、长相等个体差异有关，是决定人类疾病（尤其是多基因疾病）易感性和药物反应性差异的核心信息，同时，某些 SNP 的频率在不同民族、人群有明显差异，甚至部分 SNP 显示群体专一性。这就使得 SNP 对生物医学研究、药物开发、

医学诊断和法医学发展有重要意义。除了基因制药、诊断、生物医学研究方面的应用，SNP 图谱还有望被用来识别基因组上成千上万个附加标记，以简化 HGP 研究者绘制的很大的基因组图谱。

(一) SNP 与疾病的关系

基因在决定个体的正常表型，即形态、代谢和免疫状态等方面起着决定性的作用。通过赋予个体对疾病的易感性或抵抗力，以及影响机体与环境因素的相互作用，基因也对任何疾病的发生发展起着重要作用。

复杂的多基因疾病，如癌症、糖尿病、血管性疾病和某些精神性疾病，很难用传统寻找基因的方法来确定，因为仅仅一个基因的改变对其影响不大。随着 SNP 的不断发现和人类第三代遗传标记图的绘制，现在已有可能描绘在某一疾病发病时或发育阶段中多个基因位点甚至整个基因组的状态，从而对复杂的多基因病产生更深的认识。目前利用 SNP 可以在以下几个方面发挥作用：

(1) 确定疾病相关基因：进行简单和复杂疾病的遗传连锁分析(linkage analysis)及关联分析(association analysis)，用于疾病易感基因定位；而且其定位的精度将比微卫星标记精细得多，可直接用于指导易感基因克隆。

(2) 寻找可以带来新治疗方案的靶基因。

(3) 发现预测治疗效果的基因。

(二) SNP 与环境的关系

人们已经发现不同种族、不同人群对外界环境因素变化的反应不同，阐明这一现象的本质有赖于对 SNP 的研究。

在疾病发生、发展的过程中，个体或群体对环境致病因素的易感性起着重要作用。这种易感性的遗传基础是基因组的结构差异或(和)表达差异。对于大多数多基因疾病，一些基因的基因多态性不一定与疾病的易感性直接有关。但是疾病一旦发生，这些基因多态性对其编码蛋白质的不同影响，将会使疾病的临床表型在不同基因型的个体之间出现差异。SNP 有助于阐明这些差异。

人类的易感性也并不完全由基因决定。在环境致病因素作用下的基因表达差异起着更为重要的作用。随着 DNA 芯片在基因表达研究中进一步应用，在理论上只需很少的个体或标本就可确定环境因素对基因组表达的影响并找出易感基因。

(三) SNP 与个体化治疗

基因多态性造就了人与人之间在临床表型(clinical phenotype)上的个体差异性。在基因多态性研究的引导下，临床医师将有可能预测不同的个体在同样的致病条件下会出现什么样的病理反应和临床表现，也就是不同的临床表型，不论在疾病的预防和治疗上都能采取更具个体化的措施。

因此，从临床研究出发，根据疾病病理生理特征性改变，选择合适的候选基因分析其多态性与临床表型(症候、转归、治疗反应)之间的联系，将为临床上从基因水平认识多基因疾

病亚临床表型的本质及有选择地采取个体化治疗奠定基础。

(四) SNP 与药物的关系

同一种药物用于同一种疾病的不同病人,效果不同,造成这种差异的原因是由于药物本身在不同个体体内活化、代谢、清除方面的差异所决定的,而这种差异本质是基因差异。SNP 能充分地反映个体间的遗传差异。通过研究遗传多态性与个体对药物敏感性或耐受性的相关性,可以阐明遗传因素对药物效用的影响,从而指导医师根据个人的基因特征来选择药物及药物剂量,也即真正做到用药个体化。

将药物疗效与基因多态性结合起来,有可能找出某些特殊人群。在他们身上那些被临床试验证明是"无效"的药物仍可能发挥作用,并可从中排除那些会出现不良反应的患者。这将会大大提高药物效果,同时减少药物不良反应的发生。

目前,正在兴起的"药物基因组学"(pharmacogenomics)研究中,可通过检测 SNP 的遗传多态性标记揭示人群中不同个体对不同药物的敏感性差异的根本原因,为临床有针对性地合理用药和根据不同基因型群体对药物的反应来改进药物设计提供了理论依据。这是当前制药行业对 SNP 表现出空前兴趣的原因。

(五) SNP 与其他

SNP 也可用于法医研究中的罪犯身份的鉴别、亲子鉴定等,此外在器官移植中供体和受体间的配对选择及物种进化的研究中都将具有重要意义。

四、SNP 检测方法

SNP 本质上是 DNA 序列中单个碱基的置换,理论上任何用于检测单碱基突变或多态性的技术都可用于 SNP 的识别或检出。

目前已有多种方法可用于 SNP 检测,但不管哪一种方法,首先必须进行靶序列的扩增,然后才能进行其他检测。

传统的 SNP 检测方法是采用一些已有的成熟技术,如 DNA 测序、限制性酶切片段长度多态性(RFLP)、单链构象多态性(SSCP)、等位基因特异的寡聚核苷酸杂交(ASO)等。这些技术虽然在某种程度上能完成对 SNP 的检测,但由于它们必须通过凝胶电泳进行检测,因此,距快速、高效、自动化的目标还相差甚远。传统的 RFLP 只能检测到 SNP 的一部分,测序技术既费时费力,又不易实现自动化,而且 DNA 链的二级结构还容易造成人工假相,使测序结果出现偏差,不适宜 SNP 的检测。

目前,利用现有技术进行大规模、大范围内检测 SNP 还有一定难度。DNA 芯片可能为 SNP 的研究带来光明前景。

DNA 芯片(DNA chip)技术是近年来新开发的一种 DNA 序列变异检测工具。其主要原理是在一块小硅片上进行微阵列分析,让目标 DNA 与密集的多重寡核苷酸阵列进行杂交,进而检出 SNP 的有效方法。

目前已有多家公司开展了对芯片的研究,例如 Research Genetics 公司新近开发了 1 个

集成有 1500 个 SNP 的 DNA 芯片,它涵盖了人类基因组全部 24 条染色体,所提供的信息量至少等于或优于目前常用的 300～400 个微卫星标记的图谱,检测时只需 0.5μg 的 DNA 样品就可进行 1 次全基因组的扫描。

五、SNP 的网上资源

目前已有许多生物医学网站开辟了专门的 SNP 网页,可以很方便地在这些网站上查阅有关的 SNP 信息。国际上较重要的网站有:①dbSNP(http:// www. ncbi. nlm. nih. gov/SNP/):该网站是由美国的 NCBI 主办的。它除了可接受各地发来的 SNP 申请注册外,也向公众免费提供对 SNP 的查询。②HGBASE(http://hgbase. interactiva. de):该网站建在德国,收集基因内 SNP,研究者可通过检测出的序列查询 SNP。③MIT SNP 数据库(http://www-genome. wi. mit. edu /SNP/human/ index. html):该网站是由美国麻省理工学院建立的。它包括数千条已经定位的 SNP,可以通过指定染色体的某一区域查询 SNP。

此外,可供利用的公开 SNP 网上资源还有:

由美国国立卫生研究院(National Institute of Health, NIH)提供的主要是与癌症和肿瘤相关的候选 SNP 数据库: http://cgap. nci. nih. gov/GAI

由人类基因组组织机构(Human Genome Organization, HUGO)维持的突变数据库: http://ariel. ucs. unimelb. edu. au/cotton/mdi. htm

由美国白头研究所(Whitehead Institute for Biomedical Research Genome Institute)建立的人类 SNP 数据库: http://www-genome. wi. mit. edu/SNP/human/index. html

由华盛顿大学(Washington University)支助的按染色体位置组织的 SNP 数据库: http://www. ibc. wustl. edu/SNP

由瑞典卡尔林斯卡研究院(Karolinska Institute of Sweden)建立的 HGBase 数据库: http://hgbase. cgr. ki. se/

由国际医药与信息加工公司联合组成的 SNP 研究联盟(The SNP Consortium, TSC)建立的 SNP 数据库:http://snp. cshl. org/db/snp/map

由美国国立环境健康科学研究院(National Institute of Environmental Health Science)资助的犹他州大学 SNP 数据库: http://www. genome. utah. edu/genesnps/

(马春红)

第十七章　SEREX 方法筛选功能基因

Chapter17. Screening of Target Genes by SEREX

人类基因组计划的核心工作是 DNA 测序，然而完成人类全基因组 DNA 序列的测定只是解密人类遗传密码的基础，更重要的和更大量的工作是功能基因组的研究。因为人类最想知道的是，这些线性排列的核苷酸序列最终是如何实现其功能的。

虽然大规模 cDNA 测序和 DNA 芯片技术加快了新基因发现的速度，但由于缺乏功能方面的联系，因而有一定的局限性。因此，通过功能筛选来寻找新基因渐渐成为一个新的趋势。利用生物学功能、表现型差异、组织特异性、细胞耐药性和免疫原性等特点进行大规模功能基因的筛选的方法近些年来层出不穷。这些方法的优点是由于筛选本身是与功能联系在一起的，因此初选的基因与功能有一定联系，这样就大大缩小了与功能研究目标的距离，加快了功能研究的速度。

基于基因工程学原理的重组 cDNA 表达文库的血清学分析法(serological analysis of recombinant cDNA expression libraries ,SEREX)就是一种功能基因的筛选方法，该方法利用自身抗原或肿瘤抗原的免疫原性来筛选与自身免疫病或肿瘤相关的抗原基因，然后进一步研究这些基因与自身免疫病或肿瘤发生、发展的关系，为疾病治疗奠定基础。

本章以 SEREX 方法筛选肿瘤抗原基因为例，简单介绍基因工程学技术在功能基因筛选中的应用。

第一节　SEREX 方法原理与技术步骤

一、SEREX 方法原理

SEREX 方法的理论基础是人体的 B 淋巴细胞可以识别自身肿瘤抗原，因此肿瘤患者血清中可能含有肿瘤抗原的相应抗体。建立肿瘤 cDNA 表达文库，使肿瘤抗原浓度成百倍的增加，然后用肿瘤患者血清筛选文库。阳性克隆中插入的 cDNA 被认为可能是对应于肿瘤抗原的编码基因，进一步分析阳性克隆，鉴定出相对特异的肿瘤抗原基因。

二、SEREX 技术步骤

(一) 建库

噬菌体 λZAP 在大肠杆菌 *lacZ* 启动子下游有一个多克隆位点，外源性 cDNA 可插入该

位点，从而在感染菌或诱导的溶源体中进行表达。提取病人肿瘤组织的 RNA，分离纯化其中的 mRNA，合成 cDNA 片段。反转录时所用 Oligo(dT)引物带有 *Xho* Ⅰ酶切位点，然后两端连接 *Eco*R Ⅰ 适配子，最后再经 *Xho* Ⅰ酶酶切，这时的 cDNA 一端是 *Xho* Ⅰ黏末端，另一端是 *Eco*R Ⅰ黏末端，这样，cDNA 就可与 λZAP 噬菌体表达载体的 DNA 重组。重组子经过体外包装，可获得大约 1～2×10⁶ PFU 的噬菌体，构成 cDNA 表达文库(图 17-1)。

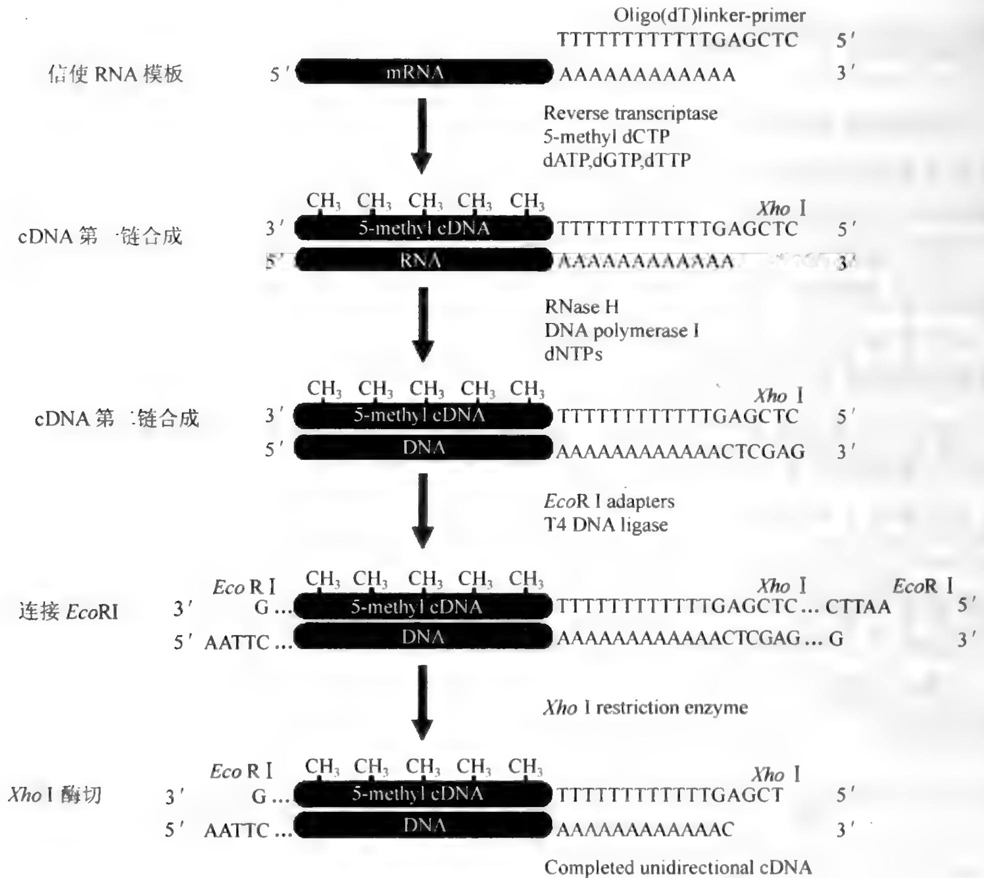


图 17-1 cDNA 合成流程

(二) 筛库

把重组噬菌体转染受体菌后，接种于固体培养基上，过夜培养后可见噬菌斑形成，噬菌斑中含有重组基因的编码产物。把噬菌斑转移至硝酸纤维素膜上，经病人血清孵育后再用碱性磷酸酶标记的羊抗人 IgG 孵育，然后通过碱性磷酸酶催化的显色反应观察病人血清与膜上的蛋白成分是否发生血清学反应，从而对所建 cDNA 文库进行免疫筛选。第一轮筛选所得阳性克隆再进行第二轮或第三轮筛选，直至得到阳性单克隆。

（三）测序

用含有阳性重组 DNA 的 λZAP 噬菌体和辅助噬菌体（单链噬菌体）共感染细菌。建库所用的 λZAP 噬菌体载体 DNA 中构建有单链噬菌体 f1 噬菌体的复制起始信号和终止信号，位于多克隆位点的两侧。在同一宿主菌中辅助噬菌体 f1 所产生的复制相关基因产物可以识别这些信号，首先在 λZAP 噬菌体 DNA 链的起始信号部位造成切口，然后合成单链 DNA 的反应越过多克隆位点（及任何插入的 cDNA 序列）而止于下游终止位点。其中一些子代分子可环化并包装入 f1 颗粒，然后从感染细胞中分泌。这个过程被称之为 λZAP 噬菌体载体 DNA 的体内剪切。

这些分泌颗粒再次感染细菌后，进入细胞的单链 DNA 转变为带有 ColE1 复制起点的双链分子，被称之为噬菌粒 pBK-CMV，可按一般质粒的方式进行复制，且可用常规方法加以扩增和纯化。然后提取质粒 DNA，酶切鉴定后对插入的 DNA 进行测序（图 17-2）。

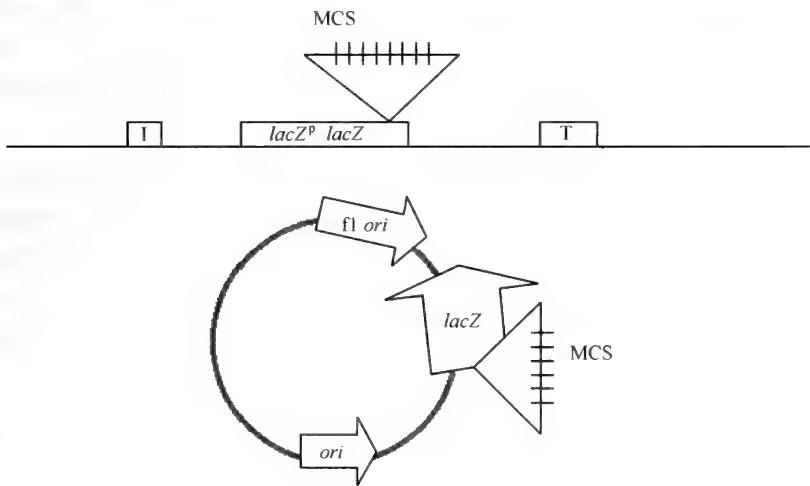


图 17-2　上：　建库所用载体　λZAP 表达载体

下：体内剪切后形成的噬菌粒 pBK-CMV

含有 *lacZ pr* 启动子和 *lacZ'* 基因；MCS 中含有 *Eco*R Ⅰ和 *Xho* Ⅰ酶切位点；含有 f1 丝状噬菌体复制起始信号 I 和终止信号 T；含有噬菌粒 pBK-CMV 序列

（四）生物信息学分析

使用 BLAST 工具把测知的 DNA 序列与 GenBank 中的数据相对照，以确定该序列是否是一新基因。

肿瘤相关基因被筛选出来之后，可通过 RT-PCR 和 Northern blot 技术进一步分析其表达谱，观察该基因在人体各组织和肿瘤组织中的表达情况，以判断其肿瘤特异性；或者进一步分析该基因的功能，阐释其与肿瘤发生发展的关系。

第二节 SEREX 方法所筛选的肿瘤相关抗原基因

目前,SEREX 方法已用于多种肿瘤的研究,包括黑色素瘤、食管癌、肺癌、星形细胞瘤、神经角质瘤、胰腺癌、前列腺癌、乳腺癌、白血病、脑膜瘤、头颈部肿瘤、胃癌、恶性间皮瘤、肝癌、肾细胞癌、结肠癌和卵巢癌等。现已鉴定出多种肿瘤相关抗原基因,这些基因大多数在 SEREX 网站(http://www.licr.org/SEREX.html)都有登陆,可被归纳为以下几类:① CT 抗原(cancer-testis antigens):一类仅出现于某些肿瘤组织及正常睾丸组织中的肿瘤抗原。②突变抗原:一类由于肿瘤组织基因突变所导致肿瘤相关抗原。③过表达抗原:在肿瘤组织中过量表达,超出维持机体自身耐受而诱发免疫应答的自身抗原。④分化抗原:高表达于特定组织肿瘤,在其相应正常组织中低表达,在其他正常组织或其他肿瘤中不表达。⑤剪切变异抗原:在肿瘤细胞中常有正常基因不同 mRNA 剪切体产生,所编码的新的变异体蛋白导致免疫应答。有一些 SEREX 方法所筛选的肿瘤相关抗原基因又被做了更进一步的研究,尤其是 CT 抗原基因,由于它们的表达主要局限于肿瘤和睾丸等生殖系统的一些组织,而其他正常组织不表达,所以是肿瘤免疫治疗的理想靶位,其中一些在多种肿瘤中广泛表达的 CT 抗原,如 NY-ESO-1 已被作为肿瘤疫苗用于临床试验。

由此可见,利用基因工程学原理设计适当功能基因筛选方法,先初步筛选出与功能有一定联系的基因,这样就大大缩小了与功能研究目标的距离,可以尽快地用于临床实践,从而提高了研究的效率。同时,由于 SEREX 方法还是一种规模筛选方法,每次筛选都可能有一批与目标功能有关的基因被发现,这也就加快了发现新功能基因的速度。可以相信,随着大规模功能筛选技术和方法的不断改进与完善,发现有功能的新基因的速度会越来越快。

(石永玉)

第十八章　基因敲除与转基因技术

Chapter18. Gene Knockout and Transgenic Technique

基因敲除(gene knockout)是近10年来在转基因技术和人工同源重组技术的基础上发展起来的高新生物技术,又称基因打靶(gene targeting)。这种技术是通过基因工程的方法将一个结构已知但功能未知的基因去除,或用其他序列相近的基因取代(又称为基因敲入,gene knockin),然后从整体观察实验动物,从而推测相应基因的功能。这种人为地把实验动物某一种有功能的基因完全缺失的技术称为基因敲除技术。该项技术是Marrio Capecchi于20世纪80年代末在犹太州(Utah)大学发展起来的,实验用的动物通常是小鼠。这种被敲除了某种功能基因的小鼠称为基因敲除小鼠(knockout mice)。

所谓转基因技术,是指利用物理、化学和生物学的方法将人工分离和修饰过的基因导入到生物体基因组中,由于导入基因的表达引起生物体性状的可遗传性修饰,并能稳定地遗传给后代,从而建立转基因种系或转基因群。通过实验导入的方法使外源基因在实验动物染色体基因组内稳定整合,并遗传给后代,这种稳定整合了外源基因的动物称为转基因动物(transgenic animals)。

第一节　基因敲除的原理

基因敲除是20世纪80年代后半期应用DNA同源重组原理发展起来的一门新技术,是同源重组技术的形象说法,即是将外源DNA与受体细胞基因组上的同源序列之间发生重组,并整合到预定位点上,使受体细胞特定的内源性基因被破坏,而不累及其他基因,从而改变细胞的遗传性状而造成其相应的功能丧失,这与早期生理学研究中常用的切除部分-观察整体-推测功能的三部曲思想相似。基因敲除不仅可中止某一基因的表达,还可引入新基因及引入定点突变,既可以用突变基因敲除其相应的正常基因,也可以用正常基因敲除相应的突变基因。

根据重组后靶基因的特征,基因敲除大致可分为两种类型:①基因破坏或剔除:由于外源序列的引入或部分取代,靶基因原有结构被破坏;②基因置换:靶基因的全部序列被新的基因或改造后的基因所取代。

第二节　基因敲除的策略

20世纪80年代初,胚胎干细胞(embryonic stem cells,ES细胞)分离和体外培养的成功奠定了基因敲除的技术基础;1985年,首次证实的哺乳动物细胞中同源重组的存在奠定了

基因敲除的理论基础；到1987年，Thompsson首次建立了完整的ES细胞基因敲除的小鼠模型。目前最常用的基因敲除靶细胞是小鼠ES细胞。基因敲除是一套组合技术，包括基因重组、细胞分离培养、转基因等，基因敲除的技术路线如下：

(1) 构建重组基因载体；

(2) 用电穿孔、显微注射等方法把重组DNA转入受体细胞核内；

(3) 用选择培养基筛选出已击中的细胞；

(4) 将击中细胞转入囊胚腔，并移植回假孕母鼠获得生殖系嵌合体小鼠，经过适当交配，获得源于ES细胞的纯系小鼠，进行形态学观察及分子生物学检测。

一、基因敲除载体的设计

DNA间发生同源重组的频率是很低的($10^{-3}\sim10^{-7}$)，所以在设计基因敲除载体时，提高同源重组发生的频率和引进选择系统是试验成功与否的关键因素。应用同源基因DNA片段构建载体，可以将同源重组频率提高20倍，随着同源臂长度的增加，重组频率也增加。

二、基因敲除载体的构建

构建载体时，首先要获得与ES细胞相同品系的基因片段作为同源片段插入载体，并在载体上插入筛选标记基因。

构建基因敲除载体的基本过程如下(图18-1)：

(1) 获得目的基因(待敲除基因)的同源片段，将此DNA片段克隆到一般的质粒载体中。

(2) 从重组质粒中切除目的基因同源片段的大部分同源DNA序列，只留部分序列在线性质粒载体的两端。

(3) 将新霉素抗性基因(*neo*)克隆到带有目的基因同源序列的线性质粒中，位于残留目的基因同源序列的中间。

(4) 在目的基因同源顺序的外侧线性化重组质粒载体，将疱疹病毒胸苷激酶(HSV-tk)基因克隆到此线性载体中。

因此，基因敲除载体由部分残留的待敲除基因的同源片段、位于其内部的*neo*基因和位于其外侧的HSV-*tk*基因共同构成。

三、基因敲除载体导入ES细胞

一般多采用显微注射法(或电转移的方法)将基因敲除载体导入ES细胞中，通过基因敲除载体上目的基因的同源顺序与染色体上的待敲除基因发生重组置换，以载体上的*neo*基因置换ES细胞基因组中的目的基因，从而得到了丧失目的基因功能的ES细胞，即基因敲除细胞。

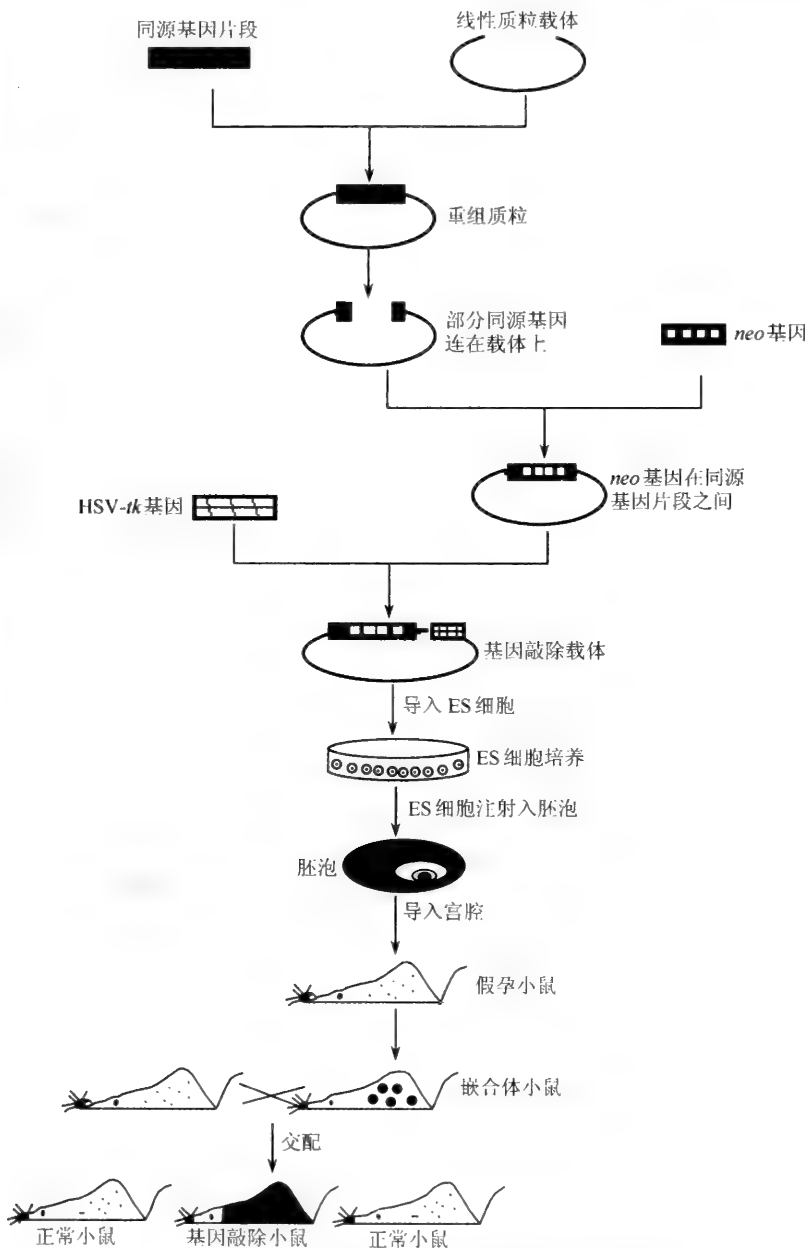


图 18-1 基因敲除载体的构建及基因敲除流程图

四、阳性细胞的筛选与鉴定

基因敲除的技术路线虽不复杂,但由于高等真核细胞内外源 DNA 与靶细胞 DNA 序列自然发生同源重组的几率非常低,约为百万分之一,要把基因敲除成功的细胞筛选出来是一件非常困难的工作。因此,同源重组的筛选和鉴定就成了基因敲除技术所要解决的关键问题。

目前已有多种筛选方法，包括正负筛选法(positive and negative selection，PNS)、标记基因的特异位点表达法及 PCR 法，其中 PNS 法适用范围广且效率较高。

(一) PNS 法

基因敲除的载体含有两种筛选标志，在同源序列的一个外显子内插有 *neo* 作为正选择的标志，在同源序列的 3′端插有不含启动子的 HSV-*tk* 基因作为负选择的标志，HSV-*tk* 基因由临近的 *neo* 基因启动子调节。在选择性培养基中需加入新霉素和致死核苷类似物(GANC)，当外源 DNA 序列未能整合到内源基因组上时，野生型的 ES 细胞中没有 *neo* 或 HSV-*tk* 基因的表达，选择性培养基中的新霉素成为致死物，细胞不能存活；当外源 DNA 随机整合至基因组上时(通常为含 *neo* 基因的同源重组序列与 HSV-*tk* 基因共同整合到内源性基因组中)，*neo* 基因的启动子同时启动同源序列中的 *neo* 基因及同源序列之外的 HSV-*tk* 基因，HSV-*tk* 基因产物可使致死核苷类似物变成致死物(图 18-2 A)；而发生同源重组时则只是同源序列(包括 *neo* 基因)整合至基因组中，HSV-*tk* 基因位于同源序列之外而不能发生同源重组，因此阳性细胞可以在含新霉素和致死核苷类似物的培养基中存活(图 18-2 B)。

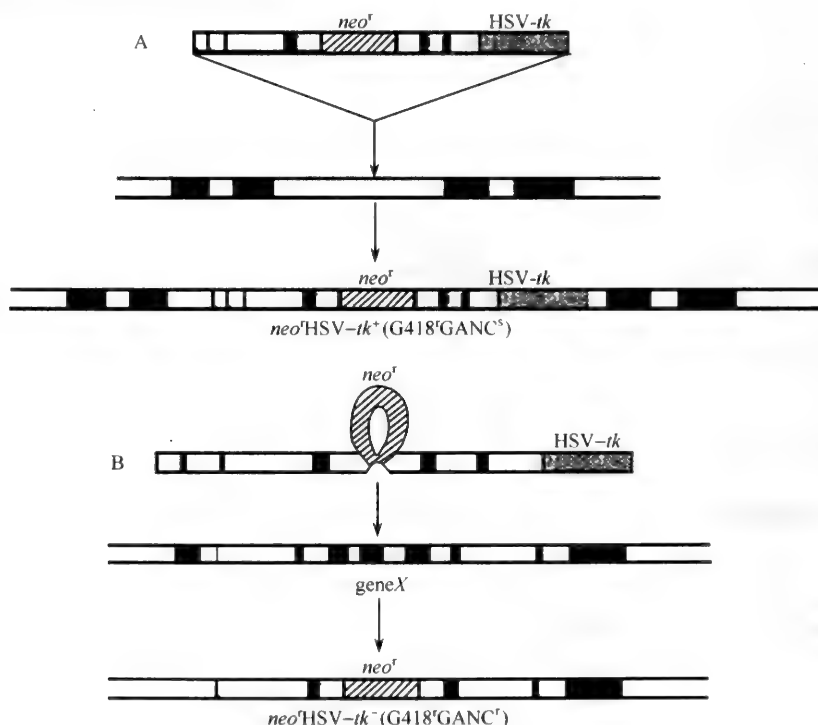


图 18-2 正负选择系统用以筛选含定位导入基因的 ES 细胞

A：随机整合；B：同源重组

neo 为新霉素抗性基因；HSV-*tk* 为疱疹病毒胸苷激酶基因；gene*X* 为 ES 细胞内源性基因(与导入基因同源)；G418 为新霉素；GANC^r 指抗致死的核苷类似物 GANC；GANC^s 指核苷类似物参与合成，细胞致死

（二）标记基因的特异位点表达法

将无启动子和起始密码子的 *neo* 基因插入靶载体中。当该靶载体转化 ES 细胞后可能出现以下几种情况：①未整合到细胞基因组而随传代消失；②随机整合，但整合位点旁侧序列中无启动子，*neo* 基因不表达；③随机整合，且整合位点附近存在其基因的启动子，启动 *neo* 基因的表达；④同源重组，基因可借助靶序列中自然存在的启动子和起始密码子而活跃表达。因此，G418 抗性细胞应具有后两种整合特性。根据它们所产生的 DNA 酶切图谱不同，可用 Southern 印迹法鉴定。此法较简单，对 ES 细胞的影响较小，但只适用于 ES 细胞中能良好表达的基因。

（三）PCR 法

PCR 法是通过选择性扩增发生同源重组的 DNA 片段来鉴定重组阳性的细胞。首先在靶载体中插入 *neo* 基因使其突变，然后合成两个引物，它们分别位于外源靶载体和与靶载体相接处的内源非同源序列上，将 G418 抗性细胞克隆分别进行 PCR 鉴定。当发生同源重组时，由于两个引物分别从相对的两个方向同时扩增，结果是 DNA 片段的拷贝数以指数增加，得到已知长度的 DNA 双链片段。而发生随机整合时，由于只有一个引物且为单方向，扩增的片段拷贝数呈线性比例，片段为单链，长度不定。PCR 法的缺点是必须对每个细胞克隆分别进行扩增，费时费力。

五、基因敲除动物的产生

将基因敲除 ES 细胞注射入胚泡中，使其与原胚泡中的细胞共同组成胚泡的内细胞团，将胚泡移植到假孕小鼠的子宫腔或输卵管中，使 ES 细胞发育成小鼠或某种组织，对后代小鼠进行筛选，可以得到基因敲除的嵌合体小鼠，然后经过杂交育种，按孟德尔遗传规律，其后代中有 1/4 的几率为纯合子，经过将雄性嵌合体小鼠与正常雌性小鼠进行交配，就可得到基因敲除的纯系小鼠，即基因敲除小鼠。

第三节　基因敲除的应用及研究进展

ES 细胞基因敲除首次使体外精细的基因操作与小鼠的整个生长发育和生命过程得到了直接的结合，为探讨高等动物基因组结构和功能提供了有效的方法，尤其对于继人类基因组计划之后正在全球范围内竞相开展的功能基因组学研究具有极为重要的意义。由基因敲除技术产生出的特殊小鼠已经对哺乳类生物学的各领域，包括发育生物学、癌生物学、免疫学、神经生物学和人类遗传学产生了极大影响。理论上，基因敲除可适用于任何能产生 ES 细胞的物种，将来可在小鼠基因敲除成熟的基础上开展其他实验动物的基因敲除工作。

一、基础研究

1. 建立人类疾病的动物模型,治疗遗传病　基因敲除小鼠是研究疾病的发生机制、分子基础及诊断治疗的重要实验材料。在基因组的特定位点引入所设计的基因突变,可模拟造成人类遗传性疾病的基因结构或基因数量的异常。1992 年成功建立了囊性纤维化病(CF)的基因敲除小鼠模型,为 CF 的基因治疗提供了良好动物模型,并于 1993 年开始临床试验。最近将 5-HT(1A)受体基因敲除建立了焦虑的小鼠模型。通过去除多余基因或修饰改造原有异常基因可以达到治疗遗传病的目的。

2. 改造动物基因型,鉴定新基因和(或)其新功能,研究发育生物学　目前对于新基因功能的研究,多采用基因敲除法在小鼠上观察该基因缺失引起的表型变化。通过分析特定基因表达产物的生物学功能,对基因结构进行修饰,在动物发生、发育的全过程中分析基因修饰与动物表型的关系,从而揭示基因的真实功能。深入研究基因敲除小鼠在胚胎发育及各个阶段的表现,可以得到详细的有关该基因在生长发育中的作用。

3. 研究基因活动的调控机制　利用基因敲除小鼠能够人为地控制基因敲除发生的时空性,可用于探讨基因活动的复杂调控机制。

二、应用研究

1. 建立特殊的遗传工程小鼠品系,改造生物、培育新的生物品种　通过基因置换可以得到能够直接产生人源性单克隆抗体的小鼠品系,用做医药研究与应用。

2. 药物筛选和新药评价体系　通过基因剔除技术可以得到特定基因缺陷的小鼠品系,用于药物有效性评价和安全性评价。

3. 研究环境诱变剂的作用规律　将所引入的 DNA 片段作为环境诱变剂作用的靶 DNA,通过对它回收后进行结构分析,研究诱变剂造成 DNA 损伤和诱发基因突变的规律。

三、研究进展

基因敲除技术在国外已是成熟技术,国内还刚刚开始。但是,国内已有几家单位在这一领域开展了一些工作。例如,利用基因敲除系统,研究了 *Smad*3 基因的功能,鉴定了 3 个与骨骼发育有关的新基因,探讨了肿瘤坏死因子受体Ⅱ在郎格罕细胞迁移中的作用,探讨了 CD45 蛋白酪氨酸磷酸酶与小鼠表皮 γδT 细胞发育的关系,研究了极低密度脂蛋白和中间密度脂蛋白与动脉硬化的关系等等。近年来,又产生了条件性基因敲除技术,即对某一特定细胞类型或细胞发育的特定阶段某一特定基因的敲除,这项新技术在基础理论研究及实际应用中都将有着极广阔的应用前景。另外,RNA 干预技术(RNA interference, RNAi)在线虫的体内外试验中都能达到基因敲除的目的,对于哺乳动物,RNAi 能在体外培养的细胞中达到基因敲除的效果。

第四节　转基因技术

1974 年，Jaenisch 应用显微注射法，在世界上首次获得了 SV40 病毒 DNA 转基因的小鼠。1982 年，Palmiter 等将克隆的生长激素基因用显微注射法（microinjection）直接导入小鼠受精卵细胞核内，所得转基因小鼠的肝、肌、心等组织都能产生生长激素，小鼠比原个体大几倍，称为“巨鼠”，使人们意识到转基因技术的巨大潜力及其在遗传育种方面的划时代意义，将分子、细胞、整体水平统一起来。

一、转基因技术的基本原理

转基因技术的基本原理是将目的基因（或基因组片段）用显微注射等方法注入实验动物的受精卵或 ES 细胞，使目的基因整合到基因组中，然后将此受精卵或 ES 细胞再植入受体动物的输卵管或子宫中，使其发育成携带有外源基因的转基因动物，人们可以通过分析转基因和动物表型的关系，揭示外源基因的功能，也可通过转入外源基因培育优良的动物品种。

二、转基因技术的策略

传统的制备转基因动物的方法是将外源基因导入胚胎中，在胚胎细胞染色体上整合，成为细胞基因组的组成部分，伴随胚胎发育，使整个机体的细胞染色体上均带有外源基因，这一基因的正确表达和稳定遗传标志着转基因动物个体的建成。

常用的制备转基因动物的步骤是：①获得和改建目的基因；②将目的基因向生殖细胞高效转移；③受精卵或胚胎组织在合适的环境中发育；④筛选、鉴定稳定的细胞系。生产转基因动物的关键主要是对外源基因的分离、重组、载体构建，导入外源基因胚胎的培养、移植，以及获得新生动物的整合和表达的鉴定技术等，其中最关键的一环是外源基因的导入。

目前使用的制备转基因动物的技术有显微注射法、反转录病毒感染法、胚胎干细胞法和体细胞克隆技术等。

（一）显微注射法

制备有良好表达活性的目的基因，需用显微注射装置将目的基因导入动物的受精卵雄性原核，将含有外源基因的受精卵移植入假孕动物的输卵管，经过妊娠、分娩，使之发育成正常的幼仔，再用分子杂交或 PCR 法，筛选出有外源基因整合的幼仔，得到转基因个体（图 18-3）。用这种方法生产的动物约有十分之一是整合外源基因的转基因动物，是目前将外源基因导入胚胎应用最广的方法。

显微注射法产生最早、使用最广，优点是外源 DNA 整合率高，可导入的目的基因长（可达 100 kb），不需要载体，直接转移目的基因，可以直接获得纯系动物，使得实验周期明显缩短。缺点是操作复杂、设备昂贵、不易推广、导入目的基因的整合位点及整合的拷贝数无法控制，易引起整合部位的插入突变而妨碍基因表达，或造成胚胎发育障碍，重则致死。

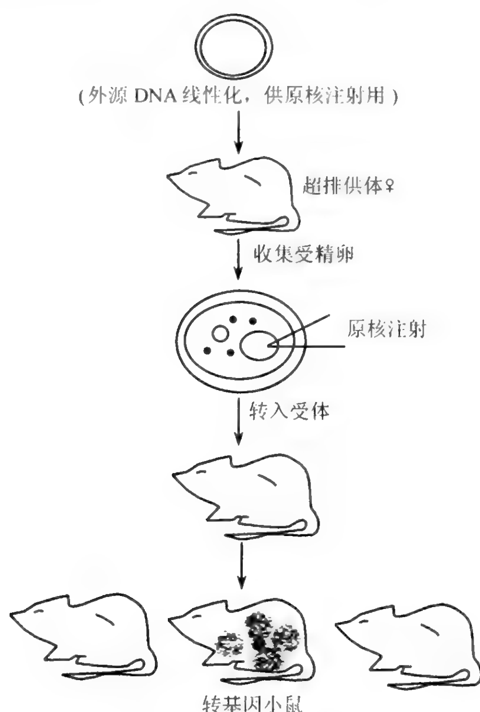


图 18-3 利用显微注射技术建立转基因小鼠的技术路线

(二) 反转录病毒载体法

将目的基因重组到反转录病毒载体上,制成高浓度的病毒颗粒,人为地感染着床前或着床后的胚胎,也可以直接将胚胎与能释放反转录病毒的单层培养细胞共孵育以达到感染的目的,通过病毒将外源目的基因插入整合到宿主基因组 DNA 中,技术路线见图 18-4。

反转录病毒载体法的优点是不需要重排,不破坏目的基因,不易发生大的突变,易分析插入位点,整合率高,整合位点单一,多为单个拷贝。缺点是需要构建带有转基因的反转录病毒,且容量有一定限制,所得的转基因个体嵌合性很高,需要进行广泛的杂交筛选,实验周期长,可能出现转入病毒基因的复制、表达等。

(三) ES 细胞介导法

ES 细胞来源于胚胎囊胚期内细胞团(inner cell masses),是多潜能细胞,移入胚胎囊胚腔时,可随胚胎发育而形成嵌合体,参与各组织器官(包括生殖腺在内)的分化,同时又具有体外培养细胞的所有特征,因此是构建转基因动物、特别是基因敲除动物时将基因导入胚胎的桥梁(详见基因敲除部分),技术路线见图 18-5。

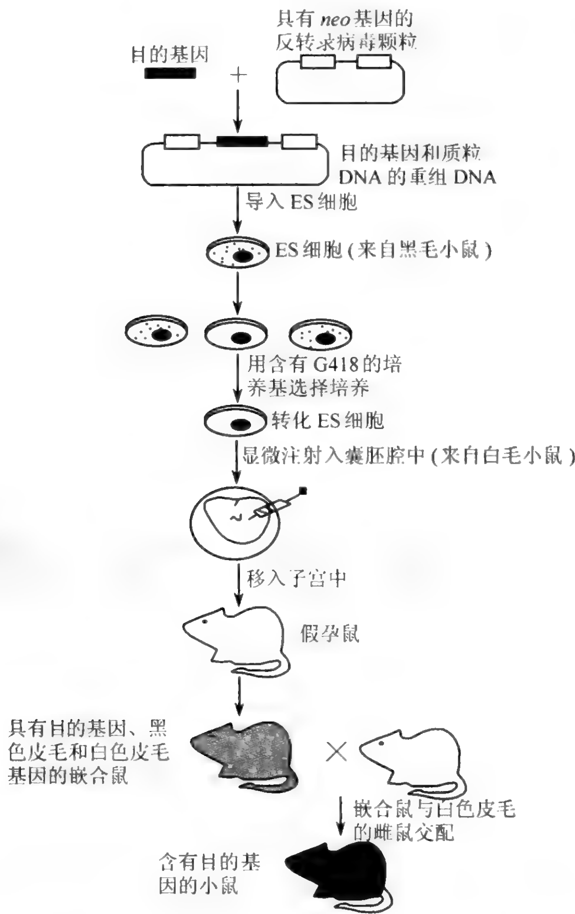


图 18-4 反转录病毒载体法建立转基因小鼠的技术路线

ES 细胞介导法是 1989 年意大利学者首先报道的，他用精子作为载体转移目的基因，成功地获得了纯系转基因小鼠。该方法优点是稳定传代的细胞系能用于各种目的基因的转移，可在植入前筛选出合适的细胞，能够得到很多遗传上相同的转基因动物。缺点是 ES 细胞系的建立和培养技术还不完善，需经嵌合体途径（许多嵌合体转基因动物的生殖细胞内不含有转基因），且实验周期长。

（四）体细胞基因转移和克隆

先在体外培养的哺乳动物体细胞中进行基因导入，筛选获得带有转基因的细胞。然后，将带有转基因的体细胞移植到去掉细胞核的卵细胞中，生产重构胚胎。重构胚胎移植到待孕母体中，获得转基因动物克隆个体，产生的仔畜全部（100%）是转基因动物。

体细胞克隆技术物种适用面广，实验周期短，效率高，能够培育出与供体基因组完全相同的转基因动物，在基础研究和应用研究方面具有重大意义。如第一例体细胞克隆动物“多利羊”的诞生引起了世界性轰动。但是，在基础理论和实验技术上体细胞克隆技术尚需进一步完善。

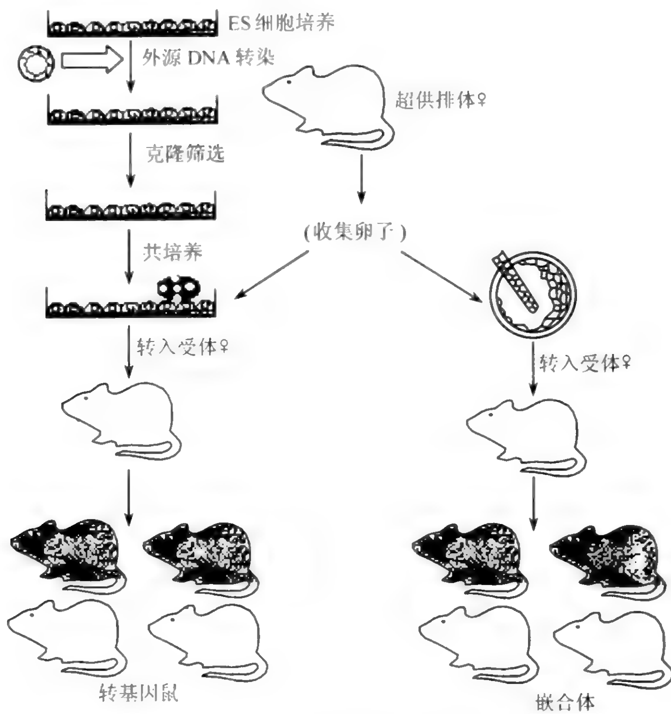


图 18-5 ES 细胞介导法建立转基因小鼠的技术路线

(五) 精子介导法

将成熟的精子细胞(通常是灭能后,也就是破坏细胞膜)与外源基因载体 DNA 进行共培养,使精子有能力携带外源 DNA 进入卵细胞中,受精并使外源 DNA 进入卵中。另外,也可通过电穿孔、脂质体转染法将载体 DNA 导入精子。将受精卵移入假孕母体输卵管中,继续发育获得转基因动物个体。

精子介导法的优点是简单、方便,依靠生理受精过程,避免了对原核的损伤,动物育种不经过嵌合体,实验周期短,对于大家畜的转基因研究具有重要意义。缺点是目的基因随机整合,无法早期验证遗传修饰事件。

三、转基因技术的应用

自 1974 年美国学者 Jaenisch 首次应用显微注射法获得转基因小鼠以来,转基因动物已用于疾病动物模型的建立、药用蛋白质的生产、转基因家畜的生产等方面。目前转基因技术已广泛应用于许多领域,如在免疫学、病毒学、代谢性疾病、神经生理学等方面的应用,在遗传育种上也取得了突破性进展。

(一) 转基因植物的研究与产业化

转基因技术用于育种，不仅可以加快改良遗传性状的进程，使选择的效率提高，改良的机会更多，而且不受有性繁殖的局限。与传统的孟德尔遗传规律育种比较，转基因技术显示出其优越性和更大的潜力，目前提高光合作用、扩大固氮能力、提高营养价值、抗虫、抗病、抗旱等转基因植物都在研究之中。将人的基因转入植物还可能获得医学上具有治疗用途的药物，例如将人抗体基因转入烟草，从烟叶中就能提取人的抗体蛋白，这为医药产业开辟了一个崭新的制药途径。

(二) 转基因动物的研究与产业化

1. 建立疾病动物模型　利用转基因技术可建立敏感动物品系及产生与人类疾病相同的疾病动物模型。转基因动物作为疾病模型可以代替传统的动物模型进行药物筛选，并且准确、经济、试验次数少、显著缩短试验时间。目前已建立的疾病模型有：癌症、动脉粥样硬化、镰状细胞性贫血、囊状纤维化、红细胞增多症、免疫缺陷、自发性高血压等。

2. 研究疾病的发病机制　利用转基因动物还可用于疾病发病机制的研究，对于人们认识、预防疾病和测试新的治疗方法提供了有力手段。例如用导入各种癌基因、致瘤病毒基因或其调控序列等的转基因小鼠，可以观察肿瘤发生的历程和影响因素；用肝炎病毒基因的转基因动物可以研究肝炎病毒基因在肝炎中的作用，利用导入各种细胞因子基因、免疫功能基因或特定核酸序列的转基因动物，可以从整体水平研究细胞因子、免疫调控、基因表达调控等问题。

3. 用于器官移植　随着器官移植手术的普及，供体缺乏成为困扰器官移植界的世界性难题。就某些器官来说，转基因动物的器官用于人体移植将有重大价值。为解决供体来源少、多老化、质量不佳等问题，人们设想将猪等动物器官作供体，但异种动物的器官进入人体内，会发生排斥反应，使移植无效。研究发现人体衰败加速因子(DAF)的基因可抑制这种排斥，若将人的 *DAF* 转移到猪身上，再将猪的器官移植到人体，会大大减弱排斥作用，此项技术称为转基因技术移植，世界上首例转基因移植实施于 1995 年。

4. 生产蛋白质药物　用转基因动物能获得治疗人类疾病的重要蛋白质，“生物反应器”的概念也由此产生。例如导入了凝血因子Ⅸ基因的转基因绵羊分泌的乳汁中含有丰富的凝血因子Ⅸ，能有效地用于血友病的治疗。现已较好地建立了动物乳腺生物反应器和体细胞克隆技术平台，目前已在下列动物的乳汁中生产出一些人类蛋白质药物：①牛：抗凝血酶、纤维蛋白原、人血白蛋白、胶原蛋白、乳铁蛋白、糖基转移酶、蛋白 C 等；②山羊：抗凝血酶原、抗胰蛋白酶、人血白蛋白、组织纤溶原激活因子、单克隆抗体等；③绵羊：抗胰蛋白酶、凝血因子Ⅸ、蛋白 C；④猪：凝血因子Ⅸ、蛋白 C、纤维蛋白原、人血红蛋白。

(三) 研究进展

从整体水平看，我国在转基因作物研究技术方面的进展与国际上基本同步，在发展中国家中居领先地位。目前我国有多种转基因植物被批准进入商品化生产，包括我国自己培育的耐储存番茄(1997)、抗虫棉(1997)、观赏植物矮牵牛(1997)、抗病毒甜椒(1998)、抗病毒番

茄(1998)。但是近年来,在国际上(尤其是欧洲)对转基因作物的安全性出现了较大的争议,这些争议在一定程度上影响了转基因作物的研究、开发和产业化的进程。

近年来转基因动物技术又有新的发展:ES 细胞导入与目的基因同源的序列,则在体内可以经同源重组使目的基因发生突变,这样成长起来的动物就有目的基因的缺陷,即基因打靶技术。用基因打靶可以在整体水平上研究基因的功能,并能制造出遗传缺陷的疾病模型。

研究功能基因组学(functional genomics)是 21 世纪全人类的一项既伟大又艰巨的工程和任务,既能造福于人类,也可能危害人类,甚至带来灾难。转基因技术可以用于生产生化武器,如果造成有毒基因的扩散,可以使某个区域内的人群丧失对某种特定病毒的免疫力,从而发生大规模的传染病。另外,在基因治疗中,由于目前对人类基因图谱并未完全破译,也可能会发生一些副作用。最可怕的是基因造人。

(高立芬)

第十九章 DNA 疫 苗

Chapter19. DNA Vaccines

第一节 概 述

一、定 义

DNA 疫苗,是指用含有编码抗原蛋白的基因或重组质粒 DNA 转染或者注射到动物细胞内,使之持续表达出有生物活性的物质,从而诱导出细胞免疫(CTL)和体液免疫(HI)效应的生物工程制品。简言之,DNA 疫苗是用编码病原体或某特定抗原蛋白组分的 DNA(或基因)诱生免疫应答的生物工程制剂。

用 DNA 疫苗直接免疫,叫做基因免疫(gene immune)或者 DNA 免疫(DNA immune)。它所诱发的不是机体对核酸分子本身的免疫反应,而是通过疫苗基因 DNA 所表达的蛋白质抗原诱导免疫系统而产生的免疫应答。

根据使用目的把 DNA 疫苗可分为两类:一类是预防性 DNA 疫苗,主要用于传染病的预防;另一类是治疗性 DNA 疫苗,多应用于肿瘤等治疗。目前,有些 DNA 疫苗兼有预防和治疗作用。大多数 DNA 疫苗现时仍处于临床前实验阶段,尚未进入临床应用。

二、构建 DNA 疫苗的基本条件

构建 DNA 疫苗的基本条件有两个:目的基因和质粒表达载体。

(一) 目的抗原编码基因

获得目的抗原基因有许多方法,原则上应选择核心蛋白基因的保守 DNA 序列,近几年采用表达文库免疫技术(expression-library immunization, ELI)可获得免疫活性基因。ELI 是根据病原体的所有抗原均由其 DNA 编码的原理,先把病原体基因文库中的病原体 DNA 片段插入到特定的质粒载体中,再利用基因免疫的方法筛选病原体基因组中具有免疫保护功能的基因片段,ELI 技术提供了一种在众多已知或未知病原体基因中获得免疫活性基因的系统而普遍有效的方法,它对于具有庞大基因组的寄生虫、细菌等基因疫苗的发展更有意义。

(二) 质粒表达载体

选择质粒表达载体的原则是既能长期、高效表达,又要免疫原性好,因此选择表达载体

时必须考虑下列组成元件:复制子、启动子、免疫激活增强序列及其增强子、终止子等。根据需要可选择细菌载体(如伤寒沙门菌 Ty21a、卡介苗)或者病毒载体(如痘苗病毒、腺病毒、单纯疱疹、HBVsAg 载体等)。目前从免疫效果看,基因免疫的载体用反转录病毒载体仍不如细菌质粒载体,如 pUC19 质粒及其衍生质粒。但是从抗病毒、抗肿瘤、防治恶性疾患的长期战略目标出发,用病毒类载体可能是有效手段之一。

1. 具有载体的复制基因　为保证疫苗 DNA 的有效扩增,可利用原核细胞如大肠杆菌作为受体菌来制备 DNA 疫苗,要求载体中必须含有大肠杆菌的 *ColEI* 或 *pMBI* 基因,才能在大肠杆菌中扩增 DNA。为确保 DNA 疫苗的安全性,防止病毒 DNA 在宿主细胞内复制或者整合到染色体上,必须去除载体中的病毒复制关键基因以及与基因整合有关的位点。例如 pcDNA3.1 载体因含有病毒的复制基因和基因整合位点而不十分安全,故多用于动物实验(图 19-1);而 pVAX1 质粒表达载体则去除了上述相关基因,有良好的生物安全性,可用作临床试验。

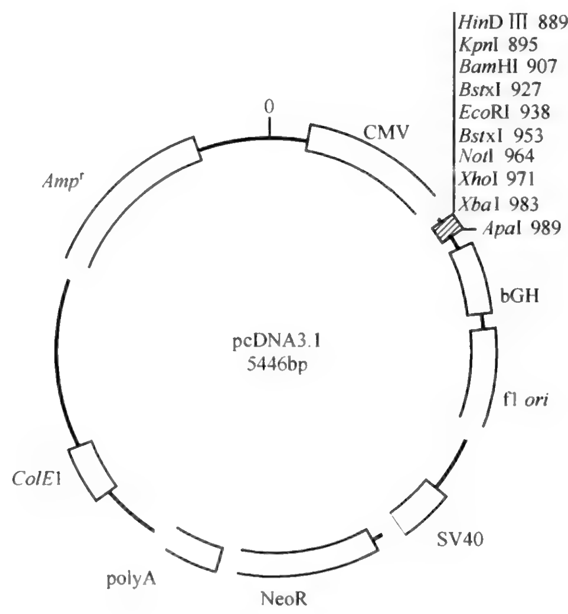


图 19-1　pcDNA3.1 结构示意图

2. 具有真核细胞的启动子　由于 DNA 疫苗是用在哺乳动物细胞内表达目的基因,必须使用能在真核细胞启动 DNA 转录的启动子,病毒的启动子具有这种作用。应用最广的病毒启动子是巨细胞病毒(CMV)、猿猴病毒(SV40)和呼吸道合胞病毒(RSV)的启动子。这些启动子均能在真核细胞中表达不同的真核和原核细胞的基因产物,其中 CMV 启动子的表达效率高,最常用。

3. 未甲基化 CpG 核苷酸序列——极好的免疫佐剂　未甲基化 CpG 核苷酸序列源自细菌 DNA 中一段由 6 个核苷酸分子组成的特异序列—X_1X_2 CpGY_1Y_2,其核心是未甲基化的 CpG 双核苷酸序列,其 5′端为两个嘌呤(X_1、X_2),3′端是两个嘧啶(Y_1、Y_2)。这种未甲基化序列在细菌中发生率很高,约为 1/16。但在真核细胞 DNA 中此序列的发生率很低,仅是细菌的 1%,这或

许是目前用细菌质粒作 DNA 疫苗载体的实验效果优于病毒载体的因素之一。

未甲基化 CpG 核苷酸序列因有激活多种免疫细胞的免疫佐剂作用，而称为免疫刺激 DNA 序列(ISS, immunostimulatory DNA sequence)。这种具有免疫增强作用的寡核苷基序可因种属不同、基因 DNA 中所含 ISS 序列和数目的不同，其免疫效果也有明显差别，因此精心设计、合成及控制不同 DNA 疫苗中的最佳 ISS，让机体获得最佳免疫反应与有效的保护作用是至关重要的。

ISS 的作用特点是：

(1) 作用细胞的广泛性：ISS 能诱导 B 细胞增殖和分泌 IgM、IL-6，诱导 $CD4^+$ T 分泌 IFN-γ；产生 T_H1 型细胞因子 IL-12、IFN-α、IFN-β、TNF-α 等，诱导 NK 细胞活化，高效地杀伤靶细胞；同时还能增强 Mφ 摄取、递呈 DNA 活性。总之，非甲基化 CpG 所诱导的免疫应答是以 T_H1 为主，这对于抗肿瘤、抗病毒疫苗的研制应用有实际意义。

(2) 作用强：ISS 佐剂作用显著，诱导的抗体比不含 ISS 序列的 DNA 高 9 倍。

(3) 不需再外加佐剂：它本身是极好的佐剂，与主分子 DNA(基因或某 DNA 片段)表现一定的剂量依赖关系。

4. 选择标记　最常用的是氨苄西林抗性基因和卡那霉素抗性基因。临床实验用 DNA 疫苗最好选用卡那霉素抗性基因，因不易产生过敏反应。

5. 其他　所用表达载体的增强子序列、最佳翻译起始序列、mRNA 终止信号、mRNA 的稳定性等都可影响基因的表达效率；基因产物的修饰、polyA 加尾信号序列和分子结构状态、分泌活性、细胞内定位等也对免疫原性也有较大影响。

三、DNA 疫苗制备的基本步骤和方法

(一) DNA 疫苗制备的基本过程

免疫原性基因→插入表达质粒载体→转化至受体细胞、扩增→体外纯化质粒 DNA→免疫接种→表达免疫原(图 19-2)。

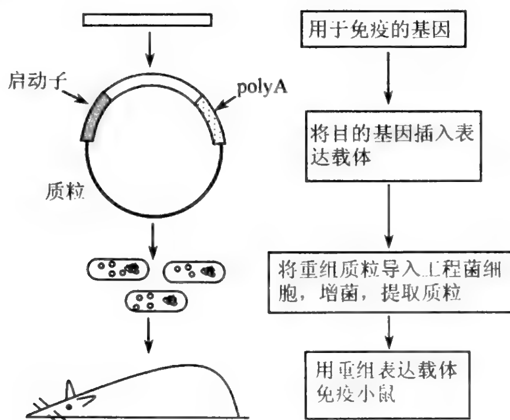


图 19-2　DNA 疫苗制备流程

(二) 主要步骤和方法

归纳如表 19-1。

表 19-1　DNA 疫苗的主要步骤、方法和注意事项

研究步骤	方法与说明
(1)基因疫苗的制备	
A.选择要表达的基因	真核或原核细胞的基因(要求基因正确); 精心设计、选择基因密码子
B.载体的选择	选择合适的启动子、CpG、选择标记和复制子
C.质粒抽提和疫苗制备	去除大肠杆菌内毒素(可用超滤方法)
(2)免疫效果的检测	
A.体外真核细胞的转染	免疫印迹证明
B.测定体内的抗体反应	ELISA 检测基因表达产物
C.实验动物模型	选用小鼠、豚鼠、灵长类动物; 检测指标用中和抗体、CTL 等
D.免疫途径	肌内、皮内、鼻内、口服和基因枪轰击; 免疫途径对 T_H1,T_H2 反应的诱导
E.免疫程序	接种剂量、间隔和次数(要制订方案)
F.攻击和免疫保护力鉴定	攻击的时间、途径和剂量; 死亡率、存活时间、体重; 脏器中细菌数和组织病理学改变
G.保护力的相关性指标	IL-2, IL-4, IL-12, IFN-γ 等细胞因子 (定性、定量、活性测定) (IgG1 和 IgG2a 水平以及两者比例)

四、常用的 DNA 疫苗质粒表达载体

(一) DNA 疫苗所用的载体

基本构成如图 19-3:

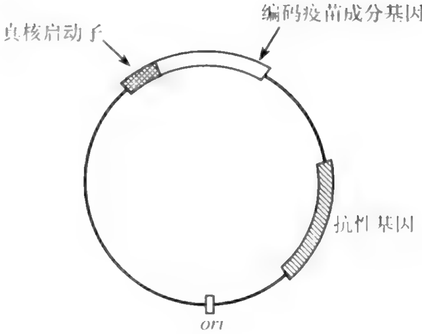


图 19-3　DNA 疫苗载体模式图

(二) 较常用的 DNA 疫苗表达载体

根据目的要求选择合适的载体质粒,较常用的 DNA 疫苗表达载体有:

1. pRC/CMV-HBs(S)(图 19-4) 即 pCMV-S,在 CMV 早期启动子控制下表达 HBV-S 抗原,专用于 DNA 免疫,不用于基因克隆。

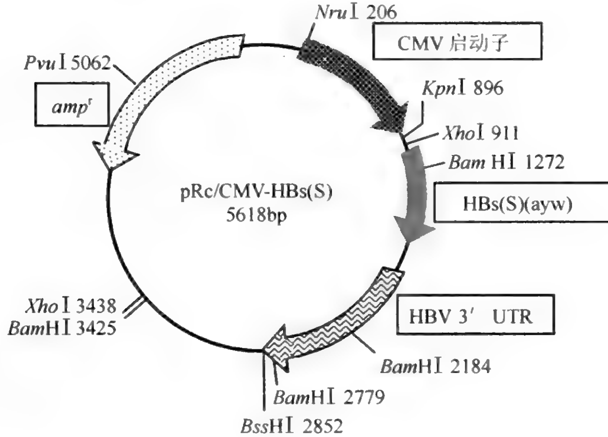


图 19-4 pPC/CMV-HBs(S)

2. 其他载体 各有特点,要选择使用,略为介绍,例如:

(1) pcDNA3.1:是克隆和表达载体,含 CMV 早期启动子、MCS 和牛生长激素腺苷酸位点,在真核细胞有 *neo* 抗性基因。

(2) pCMV-β-半乳糖酶:用于基因转染与表达,此半乳糖苷酶在 CMV 启动子控制下表达,可有较佳的抗体反应与 T 细胞效应。

(3) pCI:为 Promega 公司构建,属哺乳动物表达载体。以 pGEM(R)为骨架,含 CMV 增强子/启动子区,含有嵌合内含子(β-珠蛋白与 IgG 内含子)及 SV_4 晚期多腺苷酸信号。

(4) Vical 质粒系列:如 VR1012、VR1223、VR1332、VR1412。它们在鼠骨骼肌、肺和肿瘤细胞中表达良好。VR1012 是一个基础质粒载体,加上选择标记可构建成带有报告基因的质粒表达载体。如果加上荧光素酶基因,称为 VR1223;加上 CAT 则为 VR1332;含有 *lacZ* 报告基因,则为 VR1412。应根据需要,选用合适载体。

五、DNA 疫苗的接种

(一) 接种原则与途径

1. 接种原则 连续注射多次,间隔 2~3 周。可与蛋白质疫苗交替使用,效果更好。用多价 DNA 疫苗,优于单价 DNA 疫苗。

2. 接种途径与受体细胞

(1) 接种途径:可经鼻腔、肠道黏膜、皮肤、肌肉等途径。1999 年 9 月《自然生物技术杂

志》还报道,用皮肤涂抹法(护肤露)接种。

(2) 接种的受体细胞:DNA 疫苗免疫接种的受体细胞包括:肌细胞、角化细胞、MHC-Ⅱ阴性细胞。这些细胞使接种方便、容易吸收。

(二) 接种方式

可选择肌肉注射、皮肤转入及黏膜转入等方式。

1. 肌肉注射　肌肉注射最常用、方便有效。

用肌细胞注射 DNA 很方便,但也存在一定问题,这显然与肌细胞的特性分不开:①可长期稳定表达(肌细胞属永久性分裂细胞,分裂不活跃而不丢失 DNA,能稳定表达);②DNA 摄取效率低(可让肌细胞变性,如局部冷冻、注射麻醉药或蛇毒等,使组织处于再生状态,以提高 DNA 摄入率);③分泌不活跃(缺少共刺激分子 B_7,不易活化,MHC-Ⅰ类分子低水平表达)。④通常注射其盐溶液(单次注射 50～100μg),2～3 次,可达较好程度,以后再加强免疫。⑤基因免疫效果较好(抗体效价能维持 12～18 个月,若在 7 个月时加强免疫,抗体滴度可达原来的 10 倍;基因免疫还能诱导 T_H 细胞参与,有免疫记忆效应)。

2. 皮肤转入　可用皮内注射法和微粒子轰击法,实验研究多用后种方法。

(1) 皮内注射法(Raz 等):单次注射 3～100μg DNA,可分布于表皮、真皮细胞,诱导出 CM1、AM1。

(2) 微粒子轰击法(particle bombardment):这种技术是将疫苗 DNA 包被在重金属——金微粒子表面,用基因枪将金微粒子加速,把 DNA 分子带入靶器官、组织或单细胞内。该种方法转染效率高(几十纳克),缺点是制备 DNA 包裹颗粒方法复杂,有特殊设备的要求。

微粒子轰击法是利用火花放电、产生高压气体的电基因枪,提供转染动力。其基本方法和特点是直接转入、效率高、效果好(可诱导出 AM1、CM1)。

3. 黏膜转入　此处黏膜系指黏膜组织(包括局部黏膜、生殖道黏膜、鼻黏膜等)。

黏膜转入的方法,可用基因枪或注射法。多用于动物试验,人类应用少。

(三) DNA 免疫与加强免疫的有效匹配

DNA 免疫接种要按设计方案进行。一般接种 3 次,间隔 2～4 周,小动物(鼠类)每次 100μg,大动物或人每次 500～1000μg。

DNA 免疫接种,要注意初次免疫、强化免疫和抗原的特性。应遵循先用 DNA 疫苗,后用其他技术生产的同类疫苗,例如重组的 DNA 痘苗病毒疫苗,可诱导出 CTL 效应 ,具有感染能力和抗病毒作用。如果只用相应蛋白质、多肽抗原和佐剂,仅能增高抗体水平。

第二节　DNA 疫苗的免疫原理与应用

一、DNA 疫苗的免疫原理

DNA 疫苗免疫的确切原理,尚无定论。概括地理解 DNA 疫苗的综合作用:一是通过诱导表达的蛋白质等因子的作用;二是未甲基化 CpG 刺激顺序(ISS)的刺激诱导作用 ——即免疫佐剂作用(图 19-5)。

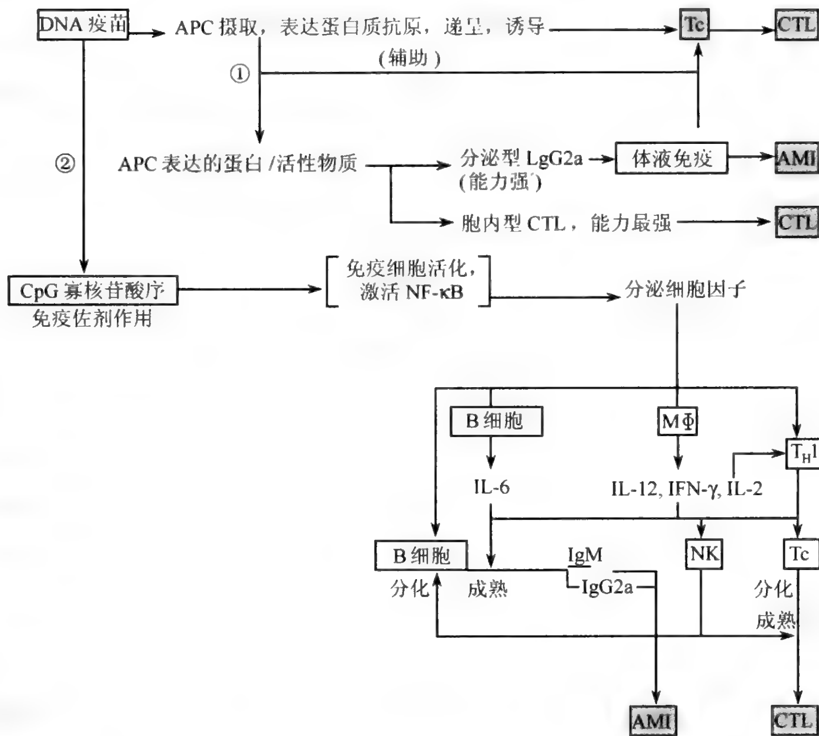


图 19-5 DNA 疫苗免疫的初步原理

DNA 疫苗的免疫效果与许多因素有关，在加强免疫时抗原的性状尤其显著。当初次免疫用 DNA 疫苗，加强免疫时用细胞内抗原，可提高 CTL 强度 10～30 倍；加强免疫若用完整的可溶性抗原，则以体液免疫为主。

二、DNA 疫苗的评价与优化

（一）主要优点

与传统疫苗、生物工程疫苗相比，DNA 疫苗有许多优点：①安全性好（无感染的危险），稳定性好（室温不易变性，活性较稳定，易于保存运输）。②具有预防、治疗的双重功能。③能诱发针对天然蛋白质表位的抗体（体液免疫）和特异性 CTL 免疫反应。④组织细胞能主动表达抗原，表达水平低，免疫反应持续时间长（长于 1 年以上）。⑤便于构建多价疫苗。⑥生产工艺技术容易控制、易操作；容易纯化，生产成本低。⑦能快速筛选有免疫保护效果的基因。⑧可定向诱导 T_H1 或 T_H2 为主的免疫反应。

（二）疫苗应用中的有关问题

目前，DNA 疫苗应用中的常见问题是：①接种的统一性和标准化（接种的剂量、部位、途径、次数、间隔等）。②DNA 疫苗生产的质量标准与应用标准化。③有致免疫耐受的可能

性(因长期低水平表达外源抗原,容易诱导机体免疫耐受),或者诱导抗 DNA 抗体的产生,发生 SLE 或其他自身免疫病。④基因整合:DNA 疫苗如果有基因整合,可能激活致癌基因,抑癌基因失活;或致基因突变,使表达异常。但是,整合往往发生在快速分裂的细胞,肌细胞是分裂终末细胞,因而肌肉注射接种较安全。

另外,基因免疫的实验效果,小动物体内效果好,大动物效果差,这可能与种属、进化或机体的免疫机能有关。

(三) DNA 疫苗的优化

成功的、能普遍应用的疫苗都必须经过优化、反复实验,才能真正像接种卡介苗(BCG)那样有效、实用。也只有把 DNA 疫苗分子的分子结构、功能基团和进入体内的命运、在体内对免疫分子和对机体的作用与影响等研究清楚,才能从作用机制上阐明。从目前 DNA 疫苗的研究技术和研究水平上看,优化 DNA 疫苗主要是从以下方面着手:①优化 DNA 载体:调节元件(包括:使用高表达的启动子 CMV,转录终止元件的设计和利用,多顺反子载体的使用等)。②添加甲病毒复制子以增强抗病毒作用:甲病毒寄主范围非常广,是一类通过节肢动物传播的正链 RNA 病毒。含自身 RNA 复制酶,能独立复制而不需利用宿主细胞的转录系统,其双链 RNA 能激活宿主细胞的抗病毒机制:一是激活蛋白激酶途径;二是激活 2′,5′-寡聚腺苷酸合成酶的途径,终致细胞凋亡。③优化 CpG 序列:寻求适合人体的最佳刺激序列(因 CpG 序列可刺激多种免疫细胞,产生多种细胞因子)。④抗原编码序列:不同物种中密码子的使用有偏爱性(bias),特别是终止密码子的改变对蛋白质抗原翻译有影响。⑤免疫佐剂的应用:基因疫苗本身在体内不能自我复制,所表达微量抗原蛋白而诱导的免疫反应通常是弱于活疫苗。用 GM-CSF、IL-4、IL-6 、IL-12 等细胞因子的表达质粒与疫苗 DNA 混合注射,能够提高机体免疫应答的水平。B7-1/IL-12 可协同 DNA 疫苗而增强 T_H1 样免疫应答,这在抗肿瘤免疫中更为重要。另外,有人试图注射 CpG-ODN 诱导 T_H1 型细胞因子释放,激活多种免疫细胞,以协同扩大 DNA 疫苗的免疫应答水平。

三、DNA 疫苗的应用研究

(一) 实验动物

实验动物以小鼠最常用,必需时再选用高等的哺育类动物。

(二) 目前研制的 DNA 疫苗

DNA 疫苗研究最为广泛的是抗病毒(传染病)疫苗,其次是抗胞内寄生的细菌性传染病(如结核病、寄生虫病)、自身免疫病和肿瘤的 DNA 疫苗。研制的 DNA 疫苗主要是:乙肝病毒 DNA 疫苗,丙肝病毒 DNA 疫苗,流感病毒 DNA 疫苗,HIV 病毒 DNA 疫苗,结核杆菌 DNA 疫苗,疟疾 DNA 疫苗和肿瘤 DNA 疫苗等。

(三) 临床试验和临床应用

1. 临床试验研究　1997 年底,FDA 已经批准 6 种 DNA 疫苗的第Ⅰ期或第Ⅱ期临床试

验:艾滋病、流感、单纯疱疹、乙型肝炎、疟疾和癌胚抗原 DNA 疫苗,其中疟疾 DNA 疫苗已进入第Ⅲ期临床实验。

2003 年传染性非典型性肺炎(SARS)的 DNA 疫苗正以惊人的速度研发着。用传统技术研制的 SARS 全病毒灭活疫苗已于 2004 年 5 月在志愿者中进行Ⅰ期临床试验,而预防 SARS 的 DNA 疫苗由于该病毒的急剧变异,其 DNA 疫苗距临床应用尚需些时间。乙型肝炎病毒、结核杆菌的 DNA 疫苗可能在 2004 年进入临床Ⅱ期。AIDS 病毒的基因重组 DNA 疫苗虽经 20 年的努力,至今尚未成功进入临床。

上述这些疫苗大多处于第Ⅰ、Ⅱ期临床试验,以证明其安全性、抗原性。至今未发现疫苗 DNA 分子整合到人体染色体中或诱发出自身免疫病的病例。疟疾 DNA 疫苗临床试验已经进入第Ⅲ期,2005 年可能作为商品供应。

2. 临床应用 DNA 疫苗目前主要应用在传统疫苗无法对付的病毒性传染病和某些恶性疾病,如艾滋病、癌症等疾病的预防和治疗。

(1) 病毒等传染病 DNA 疫苗:世界上第一例基因疫苗是治疗艾滋病的 HIV 疫苗,并于 1996 年批准在健康人体上Ⅰ期临床试验。1993 年,流感病毒的不同株间的保守性抗原基因疫苗诞生。随后 HIV、HBV、HCV、HTLV-1、HSV、支原体、结核杆菌、疟原虫等基因疫苗相继产生并逐步进入临床试验,可至今仍未有一种 DNA 疫苗真正进入临床应用。

(2) 肿瘤 DNA 疫苗:肿瘤 DNA 疫苗的研究目的,是通过宿主体内表达肿瘤特异性抗原以打破机体对肿瘤抗原的免疫耐受性,诱导出对肿瘤的特异性免疫应答。如用癌胚抗原基因疫苗诱导细胞毒 T 细胞的靶细胞杀伤效应。用 MHC-I 抗原基因疫苗治疗乳腺癌、纤维肉瘤也诱导出对肿瘤的细胞毒作用。

(3) 免疫调节疫苗:在应用基因疫苗的同时,注射某些细胞因子(如 GM-CSF, IL-4, IL-6, IL-12)或者辅助因子的基因疫苗,来增强或者调节免疫应答的方向、改变免疫应答的类型。如对某些自身免疫性疾病、移植排斥反应、多发性硬化症及避孕等 DNA 疫苗已有肯定的调节治疗作用。

DNA 疫苗也可分为预防性和治疗性 DNA 疫苗,进入 20 世纪 90 年代后,治疗性 DNA 疫苗呈现飞速发展,它的每一个进步都可能对人类肿瘤、艾滋病、肝癌、疟疾、SARS,自身免疫和某些非传染性疾病产生很大的影响。相信近 10 年内 DNA 治疗疫苗将会有突破性的成就。

研制开发疫苗,应根据传染病、非传染病以至恶性肿瘤的病因、流行病学、致病物质、致病的分子机制、过程与转归,不仅研制预防性疫苗,更要大力研发治疗性疫苗。由于基因疫苗制备的技术方法成熟、生产可标准化、研制周期较短,只要临床试验安全有效,就能进行商品化生产和成为常规的预防和治疗手段。近 10 多年来,世界许多国家尤其是发达国家对基因疫苗研发的资金投入很大,尽管目前直接应用的 DNA 疫苗还不多,但进展极快。在我国,虽然已经开展了 DNA 疫苗的研究工作,取得了可喜的成果,可研究的广度、深度、力度远不及发达国家和国际领先水平。相信,随着我国的经济发展和科技的进步,预防和治疗性疫苗的逐步应用,对于突发性、烈性传染病疫苗的研制与应用,也将为人类社会的进步和我国人民群众的生命安全、健康长寿做出贡献。

(曹英林 魏 然)



第四篇　基因工程实验技术

Part Ⅳ. Experiment Technology in Genetic Engineering



实验一　碱变性法提取质粒 DNA

Exp1. Extraction of Plasmid DNA using Alkylie Denaturing Method

【原理】

从大肠杆菌中提取质粒 DNA 的基本原理是根据细菌的染色质 DNA 和质粒 DNA 分子的大小、结构不同、碱基组成的差异等特点来进行分离的。常用的煮沸法、碱变性法、SDS 法均可获得较满意效果。碱变性法是最常用的质粒 DNA 提取方法。

碱变性法提取质粒 DNA 是根据细菌染色质 DNA 和质粒 DNA 分子的大小、结构及变性与复性的差异而达到分离的目的。细菌染色质 DNA 的分子大，为线性的双螺旋，而质粒 DNA 的分子小，为共价闭环超螺旋。在 pH12.0～12.6 的碱性环境中，染色质 DNA 的氢键断裂，双螺旋结构解开而变性；质粒 DNA 的大部分氢键也断裂，但共价闭环超螺旋结构的两条互补链不完全分离。当以 pH4.8 的高盐缓冲液调节 pH 至中性时，质粒 DNA 恢复到原来的状态保留在溶液中，但染色体 DNA 不能恢复而形成缠绕的网状结构，大部分 DNA 和蛋白质在 SDS 的作用下形成沉淀。通过离心，染色体 DNA 与不稳定的大分子 RNA、蛋白质-SDS 复合物等一起沉淀下来而被除去，质粒 DNA 存在于上清中，用酚、氯仿抽提可进一步纯化。

【试剂】

大肠杆菌 JM109-pBR322-HBV

STE：

0.1 mol/L	NaCl
10mmol/L	Tris-Cl(pH8.0)
1mmol/L	EDTA(pH8.0)

溶液Ⅰ：

50mmol/L	葡萄糖
25mmol/L	Tris-Cl(pH8.0)
10mmol/L	EDTA(pH8.0)

该溶液配制后，6.76×10^4Pa 消毒 15 min，4 ℃贮存。

溶液Ⅱ(pH12.6)(新鲜配制)：

0.2mol/L	NaOH
1%	SDS

溶液Ⅲ(pH4.8)，100 ml 含：

5mol/L NaAc	60ml

冰醋酸　　11.5ml
双蒸水　　28.5ml

TE(pH8.0)：
10mmol/L　　Tris-Cl(pH8.0)
1mmol/L　　EDTA(pH8.0)

溶菌酶(10mg/ml)、酚(饱和)、氯仿/异戊醇(24:1)、乙醇(冷)等。

一、质粒 DNA 的小量制备

【方法】

1. 细菌的培养及质粒扩增

(1) 取甘油保存的工程菌 JM109-pBR322-HBV，涂布含氨苄西林(Amp)的 LB 琼脂平板，37℃过夜。

(2) 挑取培养板上的单个菌落，接种到 2～5ml 含 Amp 的 LB 液体培养基中，37℃强烈摇荡(220 r/min)过夜。

2. 细菌的收集及裂解

(1) 取 1.4ml 培养液移至 1.5ml 的 Eppendorf 管中，12 000r/min，4℃，(或室温)离心 30s。

(2) 弃上清，1ml 溶液Ⅰ悬浮菌体 12 000r/min，离心 30s。

(3) 弃上清，将细菌沉淀悬浮于 100μl 冰预冷的溶液Ⅰ中，强烈振荡混匀。

(4) 加入 200μl 溶液 Ⅱ，颠倒混匀 5 次(不要强烈振荡)，放置冰浴中 3～5min。

(5) 加入 150μl 溶液Ⅲ，温和混匀 10s，冰浴内放置 3～5min。12 000r/min，4℃(或室温)，离心 5min。

3. 质粒 DNA 的分离与纯化

(1) 取上清移至 1 个新的 1.5ml 的 Eppendorf 管中。加入 1/2 体积饱和酚，1/2 体积氯仿/异戊醇(24/1)，颠倒混匀 2min，12 000r/min，4℃(或室温)，离心 5min。

(2) 取上清移至另 1 个 1.5 ml 的 Eppendorf 管中。加入 2 倍体积 100%冰乙醇，混匀，室温放置 5～30min。12 000r/min，4℃(或室温)，离心 5min。

(3) 弃上清，加入冷 70%乙醇 1ml，颠倒漂洗，12 000r/min，4℃(或室温)，离心 3min。

(4) 弃上清，将 Eppendorf 管于吸水纸上倒置 1min，室温放置 10～15min，或真空抽干 2min。加 20μl TE(pH8.0，含无 DNA 酶的 RNA 酶 20μg/ml)，溶解 DNA，短暂混匀，室温放置 30min 以消化 RNA。取 2μl 可用于电泳、内切酶酶切实验，或 -20℃贮存。

二、质粒 DNA 的大量制备

【方法】

1. 细菌的培养及质粒扩增

(1) 挑取培养板上的单个菌落，接种到 2ml 含 Amp 的 LB 液体培养基中，37℃强烈振荡

(220r/min)培养过夜，再取 0.5ml 接种至 25ml 含 Amp 的 LB 培养基中培养至 OD_{600}≈0.6。

(2) 取 24ml 培养液接种到 500ml 含 Amp 的 LB 培养基中，37℃强烈振荡 4～6h。加入氯霉素至终浓度 170μg/ml，37℃强烈振荡培养 12～16h。

2. 细菌的收集及裂解

(1) 将培养液移入离心管内，4000r/min，4℃，离心 15min，弃上清，用 100ml 冰预冷的 STE 悬浮细菌，再离心收集菌体。

(2) 将细菌悬浮于 10ml 冰预冷的溶液Ⅰ中，强烈振荡混匀，加入 1ml 溶菌酶(10mg/ml)混匀，冰浴放置 5min。

3. 质粒 DNA 的分离与纯化

(1) 加入 20ml 溶液 Ⅱ，颠倒混匀 5～7 次(不要强烈振荡)，放置 5min。

(2) 加入 15ml 冰预冷的溶液Ⅲ，温和颠倒混匀，冰浴放置 10min，12 000r/min，4℃，离心 20min。

(3) 将上清通过 4 层消毒纱布滤入 1 个新的离心管中，加入 0.6 倍体积的异丙醇混匀，室温放置 10min，12 000r/min，室温离心 15min。

(4) 小心弃上清，用 70% 乙醇溶液室温漂洗 1 次，12 000r/min 离心 5min。小心弃上清，倒置离心管在滤纸上，流净液体，或用消毒滤纸小条小心吸尽管壁上的乙醇，室温(或 37℃)放置 10～15min。

(5) 加 3ml TE(pH8.0)溶解 DNA 。进一步纯化可根据具体条件选用超速离心法、层析过柱法或 PEG 法(参见实验四 DNA 的纯化)。

【注意事项】

1. 操作时应戴手套，所用试剂与容器均需高压，以避免 DNase 污染。

2. 每一步操作中，加入溶液后均需充分混匀。

3. 碱变性时，要充分混匀使菌体完全裂解，一旦裂解(变黏稠)，应立即加入酸溶液中和。

4. 菌体裂解后，每步操作动作要轻，不要强烈振荡，以防损伤 DNA。

(宋 静)

实验二 DNA 琼脂糖凝胶电泳

Exp2. Agarose Gel Electrophoresis

电泳是分子生物学技术中分离、鉴定和提纯 DNA 的重要手段。

核酸分子是两性解离分子,在 pH 为 8.0～8.3 的电泳场中,碱基几乎不解离,磷酸全部解离,核酸分子带负电荷,向正极泳动。

带电颗粒在电泳场中的泳动可用迁移率表示,不同大小和构象的核酸分子通过一定孔径的支持物介质时,表现出不同的迁移率(即分子筛效应),从而使 DNA 分离开来。

电泳中常用溴酚蓝或二甲苯青为示踪染料指示样品的迁移过程。溴酚蓝呈蓝紫色,在 2% 琼脂糖凝胶中电泳时,迁移率接近 0.15kb 的双链线性 DNA。根据分离样品中 DNA 分子大小,参照溴酚蓝的迁移情况尚可决定是否停止电泳。

核酸需染色才能显出带型。溴化乙锭(EB)是一种荧光染料,其扁平分子可嵌入核酸双链的配对碱基之间,在紫外线激发下发出橙红色荧光。因其染色操作简单、灵敏,不影响 DNA 功能而最常使用。EB 染色一般是在凝胶中加入终浓度为 0.5μg/ml 的 EB,可在电泳过程中观察核酸的迁移情况。

本实验仅介绍 DNA 的普通琼脂糖凝胶电泳,聚丙烯酰胺凝胶电泳技术参见实验十二、十四。

【原理】

琼脂糖凝胶电泳(agarose gel electrophoresis)是利用琼脂糖凝胶的分子筛效应和赋形特性,把 DNA 分子在碱性缓冲体系的凝胶中泳动时分离开来。琼脂糖凝胶的分子筛孔径较大,适于分离 200bp～50kb 的 DNA 分子。凝胶的孔径取决于琼脂糖的浓度。

表实验 2-1 琼脂糖凝胶浓度与线形 DNA 分离范围

琼脂糖凝胶的浓度/%	线性 DNA 长度/bp
0.5	1000～30 000
0.7	800～12 000
1.0	500～10 000
1.2	400～7000
1.5	200～3000
2.0	50～2000

琼脂糖的浓度与适于分离的 DNA 大小有反平行线性关系。表实验 2-1 显示琼脂糖浓度与分离的线性 DNA 大小的对应关系,应根据所分离 DNA 分子的大小范围选择适合的凝胶浓度。

琼脂糖凝胶电泳操作简单、快速,分离范围广、实验成本低,是基因工程中最常用的技术之一。

【材料】

EB 贮存液:用双蒸水配成 10mg/ml,磁力搅拌使 EB 充分溶解,4℃ 避光保存。

琼脂糖凝胶：根据待测 DNA 分子大小，选择适当浓度凝胶，用电泳缓冲液配制。

TBE 缓冲液(5×)：

Tris 碱 54.0g

硼酸 27.5g

0.5mol/L EDTA(pH 8.0)20ml

加双蒸水至 1000ml，使用时 1:5 稀释。

上样缓冲液(6×)：

0.25%溴酚蓝

40%(W/V)蔗糖

溶于水中，4℃保存。

DNA 样品：pBR322-HBV 质粒。

【方法】

1. 制备 1%的琼脂糖凝胶：在电子天平秤上称取 1g 琼脂糖，加入 100 ml 电泳缓冲液，微波炉内煮沸 3min，冷却至 50～60℃时，加入 EB 溶液 5μl，备用。

2. 灌胶：将配制好的琼脂糖凝胶加热熔化，冷至 60℃，灌入胶槽中，胶厚 3.5～5mm，插入样品梳，待胶凝固后小心地拔掉样品梳。

3. 将胶置入电泳槽内，加入 TBE(1×)，使液面稍高出胶面 1～2mm。

4. 上样：取 10μl DNA 样品，加 2μl 上样缓冲液，混合后加到凝胶样品孔中。

5. 电泳：接通电源，加样侧接负极，另一侧接正极，调整电场强度不超过 5V/cm，电泳约 2～3h。

6. 当溴酚蓝在凝胶中迁移约 2/3 或 4/5 凝胶时，停止电泳，紫外分析仪上观察电泳后的带型。

【注意事项】

1. 琼脂糖凝胶熔化要均匀，灌胶时避免气泡。

2. 凝胶厚度 3.5～5mm。过薄则加样量小、样品孔易撕裂，过厚则紫外线不易穿透。

3. 加样时不要加进气泡，不要用吸头碰坏凝胶孔壁。

4. EB 是强诱变剂，应带手套操作，勿接触皮肤。见光易分解，应避光保存。

5. 电极不要接反。电路接通后，负极电解产气泡比正极多。

（赵 丽 曹英林）

实验三　DNA 酶切及相对分子质量测定

Exp3. Use of Restriction Endonuclease and Identification of Relative Molecular Mass

一、质粒 DNA 的小量酶切反应

【原理】

不同的Ⅱ型限制性内切酶能特异地识别双链 DNA 特定的核苷酸序列，并切割一定部位的磷酸二酯键，产生特异性的 DNA 片段。

酶切单位(1 单位)的确定：在进行酶切反应时，应根据 DNA 的量确定内切酶的使用单位。标准的酶切单位的确定方法，是以 λDNA 作为标准，能将 1μgλ DNA 在 50μl 反应体积内，37℃ 酶切 1h 完全切开，所用内切酶的量即为 1 个酶切单位。

酶切实用单位：消化 1μg 样品 DNA 所需酶的量(单位)＝(待酶切 DNA 上酶切位点数/λDNA 酶切位点数)×(λDNA 相对分子质量/待酶切 DNA 相对分子质量)。但在实际应用中，由于酶在运输保存过程中活性会部分降低，实际用量一般为计算量的 5～10 倍。

【材料】

质粒 DNA。

限制性内切酶。

10×酶切缓冲液。

无离子水。

【方法】

1. 在 1.5ml 的 Eppendorf 管内依次加入：

无离子水	11μl
10×酶切缓冲液	2μl
质粒 DNA(1μg)	5μl
限制性内切酶	2μl
总体积	20μl

2. 混匀，12 000r/min，短暂离心 30s，置 37℃ 水浴 2～3h。

3. 取 5μl，电泳观察酶切反应的程度。

4. 待酶切反应完全后，65℃ 水浴 10～15min，终止酶切反应。

【注意事项】

1. 限制性内切酶极易失活，操作应在冰浴中进行

2. 若反应条件不适合，如内切酶量太大，可导致酶的特异性切割改变，出现“星星”活性。

二、相对分子质量测定

【原理】

不同大小和构型的 DNA 在琼脂糖凝胶上电泳时，有不同的迁移率，对于线性双链 DNA 在凝胶电泳中，相对分子质量的常用对数值与泳动率成反比。据此可采用 DNA 分子长度(bp)的负对数值与泳动距离作图，使 DNA 相对分子质量的测定非常方便，也较准确。

把未知的 DNA 样品与已知相对分子质量的标准 DNA，在相同条件下电泳，用图解法或计算法可求出未知线性 DNA 的相对分子质量。

【方法】

1. DNA 的酶切，方法同前。

2. 1%普通琼脂糖凝胶的配制，方法同前。

3. 取 10μl 未知的 DNA 样品(pBR322-HBV/*Eco*RⅠ)，与 2μl 上样缓冲液(6×)混匀后，加到琼脂糖凝胶样品孔中。同时取 10μl(1μg)已知相对分子质量的标准 DNA(λDNA/*Hin*dⅢ)，加到琼脂糖凝胶另一样品孔中。

4. 电泳，2～3h 后，取下凝胶，用 5μg/ml 的 EB 溶液染色 30min。

5. 紫外线分析仪下观察结果。

6. 测量方法及计算结果：用分割规或透明尺测量出未知的 DNA 样品与已知相对分子质量的标准 DNA 的各条带距加样孔前缘的实际距离，用半对数纸作图，以泳动距离(cm)为横坐标，以碱基对(或相对分子质量)的常用对数为纵坐标，求出待测 DNA 的相对分子质量(表实验 3-1)。

【注意事项】

测相对分子质量时酶切一定要完全，否则出现不同构型 DNA，使测量不准确。

表实验 3-1　λDNA/*Hind*Ⅲ相对分子质量标准

片段序号	碱基对/bp	相对分子质量($\times 10^6$)
1	23130	15.00
2	9416	6.12
3	6557	4.26
4	4361	2.84
5	2332	1.51
6	2027	1.32
7	564	0.37
8	125	0.083

(宋　静　曹英林)

实验四　DNA 的纯化、回收及定量

EXP4. Purification, Reclamation and Identification of DNA

一、DNA 的纯化

常规方法抽提的 DNA 不经纯化，也可用于酶切、连接及克隆化，若标记探针、测序、转染哺乳动物细胞、转基因动物操作等则要进行纯化。高纯度的 DNA 不仅要求完全去除细菌染色体 DNA、RNA 及蛋白质，而且还要求选择质粒 DNA 的不同分子构型。

DNA 的纯化方法有：聚乙二醇沉淀法、精胺纯化法、玻璃粉洗脱法及商业化层析柱纯化法。现仅介绍比较实用的聚乙二醇沉淀法纯化 DNA。

聚乙二醇沉淀法纯化 DNA

本方法因操作简单、经济而最为常用。纯化的 DNA 可用于瞎菌转化、酶切、转染动物细胞，尤其对碱变性抽提的质粒，纯化效果更好。

【材料】

5mol/L LiCl，异丙醇。

TER(TE, pH8.0，含有 20μg/ml 无 DNase 的 RNase)。

1.6mol/L NaCl[含有 13%(W/V)的 PEG 8000]。

10mol/L 醋酸铵，乙醇，酚、氯仿。

【方法】

1. 将碱变性法大量抽提的质粒 DNA 溶于 TE 中，移入一新的 15ml 的离心管内，加入等体积的 5mol/L LiCl 混匀，10 000r/min，4℃，离心 10min。

2. 取上清移入另一个 15ml 离心管内，加入等体积的异丙醇混匀，室温放置 2min，10 000r/min于室温离心 10min。

3. 弃上清，70%乙醇溶液漂洗沉淀，10 000r/min，于室温离心 5min，弃乙醇并将离心管倒置于滤纸上，室温放置 5～10min。

4. 用 500μl TER 溶解沉淀，移至一个 1.5ml 的 Eppendorf 管内，室温放置 30min。

5. 加入 500μl 的 1.6mol/L NaCl(含 13%的 PEG 8000)，混匀，4℃，12 000r/min 离心 5min。

6. 弃上清，用 400μl TE 溶解沉淀，分别用酚、酚/氯仿、氯仿各抽提 1 次。

7. 上清移至另一个 1.5ml 的 Eppendorf 管内，加入 100μl 的 10mol/L 醋酸铵，2 倍体积无水乙醇，混匀，室温放置 10min，12 000r/min，4℃，离心 5min。

8. 弃上清,用200μl 70%冷乙醇溶液漂洗沉淀,12 000r/min 4℃离心2min。

9. 弃上清,将离心管倒置于滤纸上,室温放置5~10min。

10. 加入500μl TE溶解质粒DNA,测OD_{260}、OD_{280}以确定DNA的浓度及纯度。

二、DNA的回收

从含有各种不同大小的DNA混合物中,分离纯化特定相对分子质量的DNA片段,是基因工程的基本工作。实验室中最常采用的是以琼脂糖凝胶为介质将DNA分离后,再进行回收。目前常用的回收方法有DEAE膜(二乙基氨基己基纤维素)插片法、低融点琼脂糖凝胶法、电洗法、PEG法及商品化回收试剂盒等,可根据条件和习惯选用不同方法,如低融点琼脂糖凝胶法操作简便,回收率较高,适于大片段DNA的回收;DEAE膜插片法较简单,适于回收小于5kb的DNA片段,应用较普遍,而PEG法可直接从琼脂糖凝胶中纯化、回收大小不同的DNA片段。不同方法得到的DNA回收液,一般均可采用常规酚、氯仿/异戊醇抽提以去除杂质,使DNA得到纯化。因此,回收DNA也是使DNA得到纯化的过程。

本实验介绍DEAE膜插片法、PEG法及试剂盒法。

(一) DEAE膜插片法

【原理】

由Dretzen于1981年在Girvitz方法基础上改进而成。DEAE膜(diethylamenoethyl),即二乙基氨基己基纤维素膜,是一种阴离子交换纤维素,在它的纤维骨架链上,连接有阳离子交换基团,能够结合带阴电荷的DNA分子。在高盐的环境下,DNA分子又可以从DEAE膜上解离,从而回收DNA片段。

【材料】

DNA样品。

DEAE膜,眼科解剖刀、扁头镊子。

电泳缓冲液:TBE。

高盐洗脱液:TENa,20mmol/L Tris-HCl pH7.5;1mmol/L EDTA;1.5mol/L NaCl。

正丁醇,酚、氯仿,3mol/L NaAc,乙醇,TE等。

【方法】

1. DNA样品的内切酶酶切(见实验三)。

2. DNA的琼脂糖凝胶电泳分离(见实验二)。在紫外灯下观察到目的DNA完全分离,并确定位置。

3. DEAE膜的预处理

(1) 剪好DEAE膜,使其长略大于凝胶加样孔的长度,宽略高于琼脂糖凝胶的厚度。放入干净带盖的平皿中。

(2) 加入10mmol/L EDTA(pH8.0)溶液,浸泡5min,再用0.5mol/L NaOH溶液浸泡5min。

(3)取出DEAE膜,用无菌蒸馏水反复冲洗数次。若不立即用,可放于4℃冰箱备用。

4. DEAE 膜插片

(1) 紫外灯下观察到 DNA 完全分离后停止电泳,用眼科医用解剖刀在所需 DNA 带前沿切开,切口应略长于 DNA 带。

(2) 用扁头镊子取出处理好的 DEAE 膜,插入切口内,小心挤紧凝胶,使切口紧密闭合。

(3) 继续电泳,在紫外灯下观察到 DNA 全部迁移到 DEAE 膜上,停止电泳。

(4) 取出 DEAE 膜,浸于预冷双蒸水中数分钟。

5. 洗脱 DNA

(1) 取出 DEAE 膜,用干净滤纸吸干膜上的水滴。

(2) 把膜放入 1.5ml Eppendorf 管内,加入适量的高盐缓冲液(TENa)浸没纸片,37℃,2h 不断振荡。

(3) 10 000r/min,离心 5min,取上清放于另一 1.5ml Eppendorf 管中。

(4) 重复(2)~(3)步 1~2 次,直到在紫外灯下观察不到膜上有含 EB 的 DNA。将上清液合并。

6. 浓缩抽提

(1) 若所得的容积过多,可用 3× 体积的正丁醇抽提一次,以浓缩溶液并去除 EB,吸出含目的 DNA 的下层水相。

(2) 用酚/氯仿/异戊醇、氯仿/异戊醇各抽提一次,上清液加 1/10 倍体积的 3mol/L NaAc,2 倍体积冷 100% 乙醇沉淀 DNA,-20℃ 过夜。

(3) 12 000r/min,室温离心 20min,弃上清,DNA 沉淀用 70% 乙醇溶液漂洗一次,12 000 r/min,室温离心 10min,弃乙醇,真空干燥 2min,5~10μl TE 溶解 DNA。

(4) 取 0.5μl 上样电泳,并测其浓度与纯度。

【注意事项】

1. 适于回收 5kb 以下的 DNA 片段,对于 ssDNA 或 5kb 以上的 dsDNA 片段的回收率低。

2. 在整个回收过程中,不要让 DEAE 膜干燥,否则会导致 DNA 不可逆地结合在膜上而洗脱不下来,影响回收率。

(二) PEG 法回收 DNA

【原理】

PEG(polyethylenglycol,聚乙二醇)可沉淀 DNA。一定浓度的 PEG 可将核酸因相对分子质量的不同进行分离,在琼脂糖凝胶中电泳时形成清晰的 DNA 带,便于直接从琼脂糖凝胶中纯化、回收不同大小的 DNA 片段。回收的 DNA 可用于酶切及基因重组。

【材料】

DNA 样品,琼脂糖。

眼科手术刀、注射用针头,移液管。

TAE 电泳缓冲液(50 × 贮存液):242gTris;57ml 冰醋酸;100ml 0.5mol/L EDTA,pH8.0,蒸馏水定容至 1000ml。

PEG(相对分子质量为 8000,Sigma 产品)。

15%PEG/TAE 溶液的配制:15g PEG,加 TAE 至 100 ml,高压灭菌,灭活 DNase,4℃ 存

放备用。

酚/氯仿/异戊醇(25:24:1),100%、70%冷乙醇,TE 溶液。

【方法】

1. DNA 样品的内切酶酶切(见实验三)。

2. DNA 的琼脂糖凝胶电泳:0.8%～1.2%普通琼脂糖凝胶,含 EB 0.5μg/ml,采用 TAE 缓冲系统。

3. DNA 片段的回收

(1) 待电泳将 DNA 完全分离后,停止电泳。

(2) 在紫外灯下,于所需 DNA 带正前方用锐利的手术刀切一矩形槽(边缘稍大于 DNA 带),用注射用针头快速挑去切下的凝胶块,于槽内加入 15%PEG/TAE 溶液。

(3) 继续在 20～25V/cm 的电压下电泳,约 5～10min。在长波紫外灯下观察到 DNA 完全泳入 PEG/TAG 溶液中(不能泳出),停止电泳。

(4) 用移液管把含目的 DNA 的 PEG/TAE 溶液移入 1.5ml Eppendorf 管中。

4. DNA 片段的纯化(PEG 的去除)——常规酚/氯仿抽提法

(1) 上述溶液用酚/氯仿抽提一次,取上清于另一 1.5ml Eppendorf 管内,加 1/10 体积 3mol/L NaAc 和 2 倍体积的冷无水乙醇,混匀,-20℃过夜。

(2) 12 000r/min,离心 20min,倾去上清。

(3) 沉淀用 70%冷乙醇溶液漂洗,12 000r/min,离心 10min。

(4) 所得 DNA 沉淀真空干燥,TE 重溶。

(5) 取 0.5μl 上样电泳,并测其浓度与纯度。

5. 如遇到下述情况,则按如下方法:

(1) 若所得的 DNA 片段需连接于一个克隆载体上,则可从小槽中 DNA 浓度最大的地方取得,无需纯化,因 PEG 的存在可增加连接反应的效率。

(2) 若同时回收多个不同的 DNA 条带,可在其正前方挖多个矩形槽。

【注意事项】

此法适于对难以获得的微量 DNA 的制备,其回收 DNA 的大小可从 0.2～5kb,大至 50kb 的 DNA 片段,重复步骤也可回收,但需注意:

1. 电泳缓冲液一定不能高过琼脂糖凝胶表面。

2. 矩形槽的形状要规则,切下的琼脂糖凝胶要去除彻底。

3. 注入 PEG/TAE 溶液继续电泳时,需在长波紫外灯下观察 DNA 的泳动情况。长时间的紫外线照射,可造成 DNA 的损伤。

(三) 琼脂糖凝胶试剂盒回收法

目前使用较多的商品化试剂盒有上海生物工程公司出品的 UNIQ-10 柱式 DNA 胶回收试剂盒及 QIAGEN 公司的 QIAquick 凝胶回收试剂盒,均有详细的说明书,现就回收效率较高的 QIAquick 试剂盒简介如下:

【原理】

利用硅胶膜选择性吸附的特性,当含有目的 DNA 的凝胶溶液,经过装有硅胶膜的收集

柱时，使 DNA 选择性吸附于硅胶膜上，再利用洗脱液进行洗脱，收集洗脱液，即为含有目的DNA 的溶液。

【材料】

Buffer QG，Buffer PE。

Buffer EB：10mmol/L Tris-HCl，pH8.5。

异丙醇，3mol/L NaAc，pH5.0。

【方法】

1. DNA 的酶切、电泳分离，方法同前。

2. 紫外灯下，小心切下带有 DNA 的凝胶，并称重。

3. 使用<2%的琼脂糖凝胶，按每 100mg 胶重量 300μl Buffer QG 的比例，加入 Buffer QG，50℃水浴融化胶 10min，并每隔 2～3min，在旋涡混旋器上振荡 30s，使胶充分融化混匀。若溶液颜色为紫色或橙色，加入 10μl 3mol/L NaAc(pH5.0)，直到颜色与 Buffer QG 相同。

4. 若回收 DNA 片段的大小位于 500bp 与 4kb 之间，则不需加入异丙醇，否则，加入与凝胶等量的异丙醇。

5. 混匀后，移入 QIAquick 内管内，管的最大容积为 750μl，12 000r/min，4℃，离心2min，弃收集管内液体。重复此步，直到含 DNA 的胶溶液完全离心。

6. 在 QIAquick 内管内加入 0.5 ml 的 Buffer QG，室温放置 3min，12 000r/min，4℃，离心 2min，弃收集管内液体。

7. 在 QIAquick 内管内加入 0.75 ml 的 Buffer PE，室温放置 5min，12 000r/min，4℃离心 2min，弃收集管内液体。再次离心，12 000r/min，4℃，离心 2min。

8. 在 QIAquick 内管内加入 10～50μl 的 Buffer EB，室温放置 5min。取一个新的 1.5ml Eppendorf 管，将 QIAquick 内管放置其上，12 000r/min，4℃，离心 5min，Eppendorf 管内的液体即为回收纯化的目的 DNA。

9. 测定核酸的浓度及纯度。

【注意事项】

此试剂盒方法回收 DNA 纯度高，适于回收 70bp～10kb 的 DNA 片段，大于 10kb 的DNA 片段，则需选择相应的试剂盒。

三、DNA 的定量及纯度测定

做酶切、连接、转化及标记探针等，必须测定 DNA 的浓度及纯度，现介绍几种常用的测定方法。

（一）紫外分光光度计法

组成核酸分子的碱基，均具有一定的吸收紫外线特性，最大吸收波长是 260nm，吸收低谷是 230nm，根据这个特性，可用于测定核酸的浓度。

1. DNA 浓度测定　在波长 260nm 的紫外线下，1 个 OD 的光密度值相当于双链 DNA 浓度为 50μg/ml，单链 DNA 或 RNA 为 40μg/ml，单链寡聚核苷酸为 20μg/ml，可根据此计算

核酸的浓度。

【方法】

(1) 取5μl DNA样品或4μl RNA样品，加水至1ml，混匀，转入分光光度计的石英比色杯中。

(2) 分光光度计预先用1ml蒸馏水校正零点。

(3) 将样品杯放在分光光度计比色槽内。分别在260、280nm处读出光密度值(OD_{260}、OD_{280})，根据下述公式计算核酸浓度：

DNA样品的浓度=(OD_{260}×核酸稀释倍数×50)/1000μg /μl

RNA样品的浓度=(OD_{260}×核酸稀释倍数×40)/1000μg /μl

若DNA样品的稀释倍数为20，则OD_{260}的数值则是所测得的浓度；RNA样品的稀释倍数为25时，OD_{260}的数值则是所测得的浓度。

【注意事项】

此法仅用于测定大于0.25μg/ml的核酸溶液。

2. DNA纯度测定　分光光度法可通过测量260、280nm处的吸光值的比值，推算核酸的纯度，即OD_{260}/OD_{280}。DNA的比值为1.8，RNA的比值为2.0。若DNA的比值高于1.8，则说明尚含有RNA，可用RNA酶进行处理；若比值小于1.6，说明蛋白质或酚等杂质在其中，应再用酚/氯仿抽提，乙醇沉淀纯化。若RNA样品的比值小于2.0，也应该用酚/氯仿抽提，乙醇沉淀纯化。

紫外分光光度计法测定DNA时，有两个缺点：一是无法区别DNA/RNA的性质，以及质粒DNA或染色体DNA，二是测定用量偏多。

(二) 荧光分光光度法

DNA、RNA本身不能产生荧光，但荧光染料溴化乙锭(ethidium bromide, EB)可以嵌入核酸双链的配对碱基之间，在紫外线激发下，发出红色荧光，其荧光强度与核酸含量呈正比。使用已知的不同浓度的标准DNA做对照，可比较出待测DNA的浓度。此法可检测到1～10ng的样品，但是，所测的浓度是与标准比较而估计出来的，因而不是精确浓度。

1. 琼脂糖凝胶电泳——EB染色　该法因其简便易行、样品用量少而常使用。

不同大小的线性DNA经琼脂糖凝胶电泳分离后，采用EB染色，因其相对分子质量不同而出现电泳快慢行为的差异，泳出多条带型。用肉眼在紫外灯下观察、拍照，或在凝胶成像仪下拍照记录，与已知浓度的标准DNA相比较，便可做定性、定量。

采用此法，肉眼观察可检测分辨到0.05～0.1μg的DNA；凝胶拍照可鉴定出5～10ng的DNA。

2. 琼脂糖凝胶EB平板比较法　用含EB的适当浓度的琼脂糖凝胶灌制平板。将待测DNA与系列稀释的标准DNA溶液各1μl，分别点至凝胶平板上，于室温下放置数小时，在紫外灯下观察比较，可大致确定核酸含量。此法简便易行，但不能鉴别DNA、RNA及DNA的构型，也不能测定相应的样品的纯度。

(刘素侠)

实验五　细胞总 RNA 的抽提

Exp5. Extraction of Total RNA

【原理】

研究基因的表达和调控时，需要从组织和细胞中分离纯化 RNA。哺乳动物细胞内的 RNA 含量非常丰富，平均每个细胞内大约含有 $10^{-5}\mu g$，每克细胞大约相当于 10^8 个培养细胞，可分离出 5～10mgRNA。因此，真核细胞总 RNA 的分离相对较为容易。绝大多数 mRNA的 3′端存在 20～250 个多聚腺苷酸(poly A)尾，可以用寡聚(dT)亲和层析柱进行分离。因此，可以很方便地得到高质量的 mRNA，而只需要制备总 RNA。

RNA 酶(RNase)是导致 RNA 降解的最主要物质，在自然界中广泛存在，且非常稳定。常规的高压蒸汽灭菌方法和蛋白酶抑制剂并不能使所有的 RNase 完全失活，因此，必须使用 RNase 的强效抑制剂——DEPC(diethyl pyrocarbonate，焦碳酸二乙酯)处理 RNA 抽提过程中所用的一切仪器及试剂，具体事项如下：

1. 在玻璃烧杯中注入双蒸水，加入 DEPC，使其浓度为 0.1%(V/V，DEPC-H_2O)

2. 使用的试剂必须用 DEPC-H_2O 配制。

3. 所用塑料制品，均应用 DEPC-H_2O 浸泡，在通风柜中 37℃ 或室温下过夜，再高温高压蒸汽灭菌至少 30min，80～90℃ 高温下烘烤，干燥。

4. 若为玻璃制品或剪刀、镊子等金属制品，用 DEPC-H_2O 浸泡、高压后，150℃ 烘烤 4h 以上。

5. 注意：DEPC 为活性很强的剧毒物质，操作时必须戴手套，并在通风橱内进行。

下面介绍几种较为常用的细胞总 RNA 的制备方法。

(一) Trizol 法

【原理】

Trizol 是一种新型总 RNA 抽提试剂，内含异硫氰酸胍等物质，能迅速破碎细胞，抑制细胞释放出的核酸酶。整个反应过程应在 15～30℃ 下进行。

【材料】

Trizol。

氯仿。

异丙醇。

75% 乙醇溶液(DEPC 处理)

RNase 水溶液[含 0.1% DEPC(V/V)]

【方法】

1. 组织或细胞

(1) 组织：每 50～100mg 新鲜组织内加入 1ml Trizol 试剂，可用匀浆器进行匀浆。混匀组织的体积不应超过 Trizol 试剂体积的 10%。

(2) 单层生长细胞：吸去上清，以每 $10cm^2$ 面积的细胞加入 1ml Trizol 的比例，向培养瓶内直接加入 Trizol 试剂，溶解细胞，用吸液管吹打细胞溶解物数次，将混合物移至新的 1.5ml Eppendorf 管。

(3) 悬浮生长的细胞：1000r/min，5min，离心沉淀细胞，洗涤一次(可避免 mRNA 的降解)，加入 Trizol 试剂反复吹打溶解细胞。每 $5 \sim 10 \times 10^6$ 动物、植物、酵母菌细胞，每 1×10^7 细菌细胞，加入 1ml Trizol 试剂。

2. 优化步骤（可供选择）　对于富含蛋白质、脂肪、多聚糖或诸如肌肉、脂肪组织等细胞外成分的标本，需要增加该分离步骤。用 Trizol 试剂混匀组织后，12 000r/min 于 4℃ 离心 10min，以除去不溶性成分。离心后的上清中含有 RNA，形成的沉淀小球含有细胞外膜、多聚糖及相对分子质量高的 DNA，应去除。来自脂肪组织的标本，可有过多的脂肪，在最上层形成脂肪层，应该吸掉该层。将离心后的(即上清细胞溶解液)移入新的 Eppendorf 管，继续下一步骤。

3. 分相

(1) 将细胞溶解液在 15～30℃ 孵育 5min，使核蛋白复合物完全解离。

(2) 加入 0.2ml 的氯仿/1ml Trizol 试剂，注意盖紧盖子，用力摇动 Eppendorf 管 15s，并在 15～30℃ 孵育 2～3min。

(3) 4℃ 离心 15min，离心力不超过 12 000r/min。

(4) 离心后，混合物可分为三层：底层是苯酚-氯仿相，交界处为白色云状物，上层无色水相中含有 RNA。水相的体积大约为 60% 的 Trizol 试剂体积。

4. 沉淀 RNA

(1) 吸取上层水相至另一个新的 1.5 ml Eppendorf 管内。

(2) 加入 0.5ml 异丙醇/1ml Trizol 试剂，15～30℃ 孵育 10min。

(3) 4℃ 离心 10min，离心力不超过 12 000r/min。离心后，于 Eppendorf 管边缘和底部形成的沉淀即为所制备的 RNA。

5. 洗涤 RNA

(1) 弃上清，用至少 1ml 75% 乙醇溶液漂洗 RNA 沉淀。

(2) 混匀样品，4℃ 离心 5min，离心力不超过 7500r/min。

6. 重溶 RNA

(1) 弃上清，RNA 沉淀于空气/真空干燥 5～10 min，注意不能让 RNA 沉淀完全干燥，否则将降低其溶解度。

(2) 溶解 RNA 沉淀：加入 50μl 无 RNase 的 0.1% DEPC 水溶液，于 55～60℃ 孵育 10min。

(3) 上述样品即可用于反转录成 cDNA、Northern 印迹及杂交分析等实验，或放于 －80℃ 贮存备用。

(二) 异硫氰酸胍-氯化铯超速离心法

【原理】

异硫氰酸胍是蛋白质强变性剂,可有效抑制 RNA 酶的活性。细胞或组织裂解液经过 CsCl 为介质的密度梯度超速离心,DNA 与组织蛋白悬浮于上清液中,RNA 沉淀于管底从而分离。此法不仅能得到高质量的 RNA,而且能同时分离出细胞染色体 DNA,已成为提取哺乳动物细胞 RNA 的常规方法。但其操作复杂,流程较长,适于冷冻时间长、细胞质和细胞核不易分离的组织标本及富含 RNA 酶的组织细胞。

【材料】

GIT(异硫氰酸胍匀浆缓冲液):4.0mol/L 异硫氰酸胍,0.1mol/L Tris-Cl(pH7.5),1% β-巯基乙醇。

5.7 mol/L 氯化铯溶液: 5.7mol/L CsCl,0.01mol/L,EDTA(pH7.5)。

TE(含 0.1%SDS)。

TE(pH7.6)。

3mol/L NaAc(pH5.2)。

100%、70%乙醇溶液。

【方法】

1. 标本制备

(1) 组织标本:新鲜组织称重后,剪成约 $1cm^3$ 的组织块,不需漂洗直接加入匀浆液提取 RNA。

(2) 贴壁培养细胞:胰酶消化后离心收集,PBS 漂洗 1 次,再离心收集细胞,加入 1ml 匀浆液匀浆。

(3) 悬浮培养细胞:离心后,用 PBS 悬浮漂洗沉淀再离心收集,加入匀浆液匀浆。

上述标本若不立即提取 RNA,则可经液氮速冻后转至 -70℃ 贮存备用。

2. RNA 的提取

(1) 将处理好的样品移入匀浆器,加入 5~7 倍于样品体积的 GIT。

(2) 高速匀浆 1~2min,加入十二烷基肌氨酸钠(SLS)至终浓度 0.5%,混匀后,5000g,室温离心 10min。若是培养细胞可不离心。

(3) 在 Beckman 离心机 SW 60 离心管中,先加入 3.1ml 的 CsCl 溶液,然后在其上面小心铺加组织细胞匀浆液或步骤(2)的上清液 1.2ml,平衡离心对称管,40 000r/min,超速离心 12h。

(4) 弃上 2/3GIT 溶液,下 1/3 的溶液为含 DNA 的液体,较黏稠,可用于提取染色体 DNA,管底的胶状沉淀即是总 RNA。

3. RNA 的纯化

(1) 将离心管倒置倾斜,用 70%乙醇溶液先清洗管壁上的盐类和蛋白质,然后用 70%乙醇(室温)洗 RNA 沉淀 2min,弃乙醇,倒置离心管数分钟。

(2) 用 300μl TE 溶解 RNA 沉淀,转移至 1.5ml Eppendorf 管内,用等容积酚、酚/氯仿、氯仿各抽提 1 次。若为培养细胞,可省去酚、氯仿抽提。

(3) 上清加入 1/10 体积 3mol/L NaAc(pH5.2),2×体积无水乙醇混匀后 -70℃ 放置 30min。

(4) 12 000r/min,4℃ 离心 10min,弃上清,70% 乙醇溶液室温漂洗沉淀,2000r/min 离心 5min,弃上清,真空抽干 2min。

(5) 加入适量 DEPC 处理的双蒸水溶解 RNA,测定 OD_{260} 与 OD_{280},判断 RNA 的纯度和浓度。-70℃ 贮存。若长时间保存 RNA,可用小容积水悬浮溶解 RNA 后,再加入 3 倍体积乙醇,-70℃ 保存。使用前再离心回收 RNA,溶于 0.3mol/L NaAc (pH5.2)中。

【注意事项】

1. 本法仅适用于 1g 组织或 10^8 个细胞 RNA 的提取,超量的样品会影响 RNA 的提取产量与纯度。

2. 70% 乙醇漂洗 RNA 沉淀时,不要让 RNA 沉淀漂浮。

3. 一般情况下,步骤(4)中 RNA 沉淀为半透明的或看不见的,若沉淀为白色颗粒大团,说明 CsCl 已析出,应该反复用 70% 乙醇溶液漂洗或反复乙醇沉淀回收 RNA,以去除 CsCl。

(三) 盐酸胍——有机溶剂法

【原理】

由 Strohman 首创,经 MacDonald 于 1987 年改进而来,适用于没有超速离心机的情况下提取细胞总 RNA。盐酸胍是另一类强力的蛋白质变性剂,可溶解蛋白质,并使蛋白质二级结构消失,细胞结构降解,核蛋白迅速与核酸分离,因而盐酸胍又称为解偶剂。利用盐酸胍抑制 RNA 酶活性,组织细胞匀浆裂解后,有机溶剂抽提去除蛋白质,通过选择性沉淀使 RNA 分离。本法提取的 RNA 质量较好,但操作过程繁杂费时。

【材料】

盐酸胍匀浆液Ⅰ:8mol/L 盐酸胍,
0.1mol/L NaAc(pH5.2),
5mmol/L 2-巯基乙醇,
0.5% 十二烷基肌酸钠(SLS)。

盐酸胍匀浆液Ⅱ:8mol/L 盐酸胍,
0.1mol/L NaAc(pH5.2),
1mmol/L 2-巯基乙醇,
20mmol/L EDTA(pH8.0)。

乙醇:100%、70%。

氯仿/正丁醇(4:1)。

4mol/L NaAc,pH7.0。

3mol/L NaAc,pH5.2。

RNA 溶解液:0.2% SDS。

0.05mol/L EDTA,pH8.0。

【方法】

1. 标本制备　样品细胞的准备同异硫氰酸胍法。

2. 抽提 RNA

(1) 将处理好的样品移入匀浆器,加入 10 倍体积盐酸胍匀浆液Ⅰ,高速匀浆 1min。

(2) 将匀浆液移入 1 个离心管,5 000r/min 室温离心 10min。

(3) 取上清移至 1 个清洁离心管中,加入 1/10 倍体积的 3mol/L NaAc(pH5.2),混匀,再加 2.5 倍体积预冷的 100%乙醇充分混匀,0℃放置至少 2h。

(4) 5 000r/min 0℃离心 10min,弃上清,室温干燥数分钟。

(5) 每个细胞样品中加入 10～15ml 盐酸胍匀浆液Ⅱ,搅拌溶解。

(6) 加入 2.5 倍体积冰预冷的 100%乙醇,立即充分混匀,－20℃放置至少 2h。

(7) 5 000r/min 0℃离心 10min,去上清,室温放置数分钟以挥发乙醇。

(8) 重复步骤(6)、(7)两次(即总共用 2.5 倍体积乙醇沉淀 3 次)。

(9) 按每克组织细胞 5ml 的比例,加 0.02mol/L EDTA(pH8.0)溶解核酸。步骤如下:先加 1/2 体积 EDTA 振荡 1～2min,3000g 离心 2min;弃上清,再加入 1/2 容积 EDTA 振荡核酸 1～2min,合并两次核酸溶解液。

3. RNA 的纯化

(1) 上述核酸溶液用等体积氯仿/正丁醇(4:1)抽提一次,5 000r/min 室温离心 10min,吸出上清至另一个清洁离心管中。

(2) 加 3 倍体积 4mol/L NaAc(pH7.0)混匀,－20℃放置 1h 以沉淀 RNA。

(3) 5 000r/min 0℃离心 20min,RNA 将于管底形成沉淀。

(4) 吸出上清液(内含 DNA,可留取抽提 DNA),用 4℃预冷的 3mol/L NaAc (pH7.0)漂洗 RNA 沉淀,20℃ 5 000r/min 离心 20min。

(5) 弃上清,按每克组织细胞/1 ml 的比例加入 0.2% SDS,0.05mol/L EDTA(pH8.0)溶解 RNA 沉淀。

(6) 加入 2 倍体积冰预冷 100%乙醇混匀,0℃放置至少 2h,5 000r/min 4℃离心 10min。弃上清,70%乙醇漂洗 RNA 沉淀,5 000r/min 4℃离心 3min,去上清,室温干燥数分钟。

(7) 用 DEPC 处理的小体积双蒸水溶解 RNA 沉淀,加入 3 倍体积 100%乙醇,－70℃保存备用。若需要回收 RNA,加 3mol/L NaAc(pH5.2)至终浓度为 0.3mol/L,混匀,12 000 r/min 4℃离心回收 RNA。

【注意事项】

1. 在 RNA 的纯化步骤(5)中,若有 SDS 沉淀析出,滴加 0.1mol/L NaOH 溶液调 pH 至 7.5。

2. 整个过程均应在无 RNase 的环境中进行。

(刘素侠　曲守方)

实验六　聚合酶链反应

Exp6. Polymerase Chain Reaction

【原理】

PCR是利用酶促合成特异DNA片段的原理，模拟DNA复制过程的核酸体外扩增技术。利用合成的两个已知序列的寡核苷酸引物，在DNA聚合酶作用下，指导两条互补链上DNA合成，经高温模板变性、低温退火、适温延伸，三个步骤的反应为一个周期，循环进行。每一周期所产生的DNA又成为下一次循环的模板，如此循环使PCR产物以指数方式增长。当扩增25～30个周期后，可使目的DNA放大100万倍。

PCR的特点：

(1) 省时(几小时完成)，自动化操作简便。

(2) 灵敏度高：可检测pg量的模板DNA。

(3) 扩增效率高[扩增倍数为$(1+X)^n$，X：扩增效率，n：循环次数]。

(4) 特异性较高(由特异性引物、碱基互补、聚合酶特性决定)。

(5) 可允许一定核苷酸的错误掺入。

一、HBV基因扩增试剂盒检测标本中的HBV DNA

【材料】

试剂盒组成(20人份)

(1) 裂解液，500μl，1管。

(2) PCR反应混合液，200μl，1管。临用时分装于0.2ml Eppendorf管。

(3) *Taq*聚合酶存在于溴酚蓝内，200μl，1管，临用时分装。

(4) HBVDNA阳性对照血清，20μl，1管。

(5) 琼脂糖1支。

(6) 双蒸水，1000μl，1管。

【方法】

1. 待检血清DNA的提取　裂解HBV颗粒：取患者血清40μl，加裂解液20μl，混匀后，100℃煮沸10min，12 000r/min，离心20min，吸上清10μl，为PCR扩增的模板，或存放于4℃备用。

2. pBR322-HBV转化大肠杆菌阳性克隆的鉴定　pBR322-HBV连接后转化大肠杆菌DH5α，涂布Amp平板，此平板上生长的菌落有两种情况，阳性克隆与自身环化载体，可应用

PCR方法鉴定。挑取单个克隆，加入 20μl 裂解液，100℃煮沸 10min，12 000r/min 离心 20min，取上清 10μl 即为 PCR 扩增模板。

3. 扩增方法 取出已分装有 PCR 反应混合液及 *Taq* 聚合酶的 0.2ml Eppendorf 管，12 000r/min，离心 30s，每管加入处理好的模板 10μl，阳性对照可直接加入 10μl 阳性对照血清，阴性对照加入双蒸水，振荡混匀，12 000r/min，离心 30s，在 PCR 仪上按下列程序进行扩增：95℃预变性 3min 后，按下列步骤循环 35 次：①变性，95℃，30s；②退火，55℃，30s；③延伸，72℃，60s；最后于 72℃延伸 10min。

4. 扩增产物的检测 应用 2%的普通琼脂糖凝胶电泳检测：取 PCR 反应产物 10μl，直接上样电泳，20～30min 后，紫外透射仪内观察。用 DL2000 作为相对分子质量标志，观察扩增片段大小。

判断标准：若标本扩增带与阳性对照带处于同一位置，判定为标本阳性，否则为阴性。

二、RT-PCR（反转录-聚合酶链反应）

【原理】

RT-PCR 是以 RNA 为模板，体外扩增 DNA 的方法。其基本原理是，第一步以与 mRNA 3′端 polyA 互补的 oligo(dT)n 为引物，以 mRNA 为模板，在反转录酶的催化下，合成 cDNA；第二步以 cDNA 为模板，加入与 cDNA 互补的引物，在 DNA 聚合酶催化下，PCR 大量扩增目的基因。

本实验以 β-肌动蛋白（β-actin）的检测为例，介绍 RT-PCR 的过程。

肌动蛋白表达于任何生物细胞，且含量比较稳定。通过检测逆转录产物中 β-肌动蛋白的表达情况，推测反转录过程的成败，因此，β-肌动蛋白也叫做内参照。当药物不影响 β-肌动蛋白表达时，也可根据用药前、后 β-肌动蛋白的表达量，进行目的基因的半定量检测。

【材料】

反转录酶。

反转录引物：oligo(dT)n。

RNA 酶抑制剂。

细胞总 RNA。

PCR 特异性引物：浓度为 20nmol/L，

上游引物 Cx1：5′ACA CTG TGC CCA TCT ACG AGG GG 3′，

下游引物 Cx2：5′ATG AGT GAG TTG AAG GTA GTT TCG TGG AT 3′。

Taq 酶：低温保存于甘油内。

Mg^{2+}：1.5nmol/L。

DEPC-H_2O。

dNTP：10nmol/L。

【方法】

1. 以 mRNA 制备 cDNA 取一个 1.5ml Eppendorf 管，按下述依次加样：

DEPC 处理的三蒸水	9μl
oligo(dT)*n*	1μl (0.5μg)
RNA	1μl (0.1～5μg)
总体积	11μl

混匀，于 70℃ 放置 5min，迅速放冰上冷却，在 Eppendorf 管内按下述继续加样：

DEPC 处理的三蒸水	1.5μl
RNA 酶抑制剂	0.5μl
5×Buffer	4μl
10nmol/L dNTP	2μl
总体积	19μl

在旋涡混旋器上混匀 30s，12 000 r/min 离心 30s，放入 37℃ 水浴 5min，加入反转录酶(M-MLV)1μl(200μ)，于 42℃ 反应 60min，再放入 70℃ 水浴灭活反转录酶 10min，迅速置入冰水浴中冷却。此反应产物即为反转录的 cDNA，可取 1μl 测量反转录产物的浓度及纯度，其余放于 -20℃ 备用。

2. β-肌动蛋白的 PCR 扩增　取一个 0.2ml Eppendorf 管，按下述依次加入：

三蒸水	18.0μl
Cx 上	0.5μl
Cx 下	0.5μl
10×Buffer	2.5μl
dNTP	0.5μl
Mg^{2+}	1.5μl
cDNA	1.0 μl
Taq 酶	0.5 μl
总体积	25.0 μl

在旋涡混旋器上混匀 30s，12 000 r/min 离心 30s，放入 PCR 仪，按下列程序进行扩增：95℃ 预变性 3min 后，进行如下循环：①95℃，45s；②55℃，60s；③72℃，45s。共 30 个循环，再 72℃ 延伸 5min。

3. 扩增产物的琼脂糖凝胶电泳检测　制备 2% 的普通琼脂糖凝胶，取 2～10μl PCR 扩增产物，加入 2μl 6×上样缓冲液，混匀后加入加样孔内，40～80V 电压进行电泳。同时加入 DNA Maker DL2000，30～60min 后，在紫外线下观察结果，若出现 366bp DNA 带，则为阳性。在上样量一定时，可根据 DNA 带的亮度估计反转录的效果。

DL2000 电泳后的 DNA 条带：2000bp、1000bp、750bp、500bp、250bp、100bp。

有关琼脂糖凝胶电泳的详尽方法，参阅实验二。

【注意事项】

1. 血清标本　应使用一次性试管及吸管，以防污染标本。

2. 微量加样器应准确，酶加入量过大时常常造成非特异产物生成，引物量过大时易形成引物二聚体。

3. 全部操作过程应使用一次性塑料吸头，以防止交叉污染。

4. 为防止 *Taq* 酶失活，应在最后加入。反复冻融易致 *Taq* 酶失活，因此加样时最好不要让 *Taq* 酶离开冰柜。

（刘素侠）

实验七 DNA 的连接与转化

Exp7. Ligation and Transformation of DNA

一、DNA 连 接

【原理】

DNA 重组技术中的核心步骤之一是目的基因与载体之间的连接。DNA 连接本质上是一个酶促生物化学过程，即在一定条件下，由 DNA 连接酶催化目的基因与载体相邻的 5′端磷酸与 3′端羟基之间形成磷酸二酯键的过程。

相同或不同限制性内切酶产生的相同黏性末端，在降至退火温度时，能重新互补结合，在 DNA 连接酶的催化下，目的序列就与载体 DNA 链相连接。

【材料】

酶切线性化载体 DNA：pBR322/*Eco*RⅠ酶切片段。

酶切处理的目的 DNA：HBV DNA/*Eco*RⅠ。

T4-DNA 连接酶：购于华美生物工程公司，－20℃保存。

T4 Buffer(10×)：0.2mol/L Tris-HCl，pH7.6，

0.1mol/L DTT，

0.1mol/L $MgCl_2$。

ATP：10mmol/L，用灭菌双蒸水配制，－20℃保存。

【方法】

1. 取一个 1.5ml Eppendorf 管，按下述加样：

4μl	灭菌双蒸水
2μl	T4-Buffer(10×)
2μl	10mmol/L ATP
5μl	酶切处理的 pBR322 DNA
5μl	酶切处理的目的 DNA
2μl	T4-DNA 连接酶

总体积为 20μl，混匀，12 000r/min，离心 30s。

2. 置 12℃水浴，反应过夜。

3. 65℃水浴 5min，灭活 T4-连接酶。

4. 可直接用于转化试验。

5. 评估连接结果：可取 5μl 连接反应混合物，琼脂糖凝胶电泳分析，结果在载体 DNA 和插入 DNA 条带之后有若干条连续的微弱带型(由连接环化 DNA 的不同构型所致)，表示连接有望。

【注意事项】

1. T4-DNA 连接酶活性发挥需要 Mg^{2+} 辅助，并需 ATP 提供能量。

2. 连接反应温度应视所用的内切酶而定，应保证黏末端退火及酶活性、反应速率之间的平衡。

二、重组 DNA 的转化

【原理】

转化的基本原理是细菌处于 0℃、$CaCl_2$ 低渗溶液中，细菌膨胀成球形，转化混合物中的 DNA 形成抗 DNase 的羟基-钙磷酸混合物粘附于细胞表面，经 42℃ 短时间热冲击处理，促进细胞吸收 DNA 复合物，在选择性培养基平板上培养，可选出所需的转化子。

【材料】

pBR322 与 HBV 的连接产物。

LB 培养基。

LA 培养基：LB 培养基中加入终浓度为 50μg/ml 的氨苄西林(Amp)。

LA 平板：用 LB(含终浓度的 Amp 50μg/ml)配成。

100mmol/L $CaCl_2$：用双蒸水配制，高压灭菌，4℃ 保存。

JM109 受体菌(F^+，*lacZ*ΔM 缺失)

【方法】

1. 感受态细胞的制备

(1) 取甘油保存的 JM109 划线接种至 LB 琼脂板上，37℃ 培养 16～20h。

(2) 挑单菌落接种至 2ml LB 液体培养基中，37℃ 振荡，培养过夜。

(3) 取 0.5ml 培养物接种至 100ml LB 培养基中，37℃ 剧烈振荡，培养 4～6h 不动，至 OA_{600} = 0.4～0.5(细胞浓度约 5×10^7 个细胞/ml)。

(4) 将培养物置冰浴 2h，取 1.4ml 移入 1.5ml Eppendorf 管内，5000r/min，4℃，离心 5min。

(5) 弃上清，倒置 Eppendorf 管使残液流尽，将菌体重悬于冰预冷的 700μl $CaCl_2$ 溶液中，冰浴静置 30min。

(6) 5000r/min，4℃，离心 5min，弃尽上清，将菌体悬浮于冰预冷的 140μl $CaCl_2$ 溶液中，冰浴放置备用。

2. 转化

(1) 取 140μl 感受态细菌，置冰上。

(2) 加入 20μl(或 10μl)连接产物混匀，冰浴放置 30min，每隔 5min，轻轻摇动 Eppendorf 管 1 次。

(3) 42℃，热休克 1～2min，迅速放冰上，静置 1min。

(4) 加入 1ml LB 液体培养基，轻轻摇匀，37℃ 振荡培养 1～2h，使细菌恢复正常。

(5) 取 100～200μl 菌液涂布于 LA 平板上，37℃ 培养 12～24h。

(6) 进一步筛选鉴定。

【注意事项】

1. 试剂纯度　试剂纯度影响转化效率，尤其是 $CaCl_2$，甚至同一厂家出品的不同批号产品，转化率也有一定区别，最好避光储存于 4℃ 冰箱中。

2. 细菌的生长状态　应收获对数期或对数生长前期的细菌用于感受态的制备。

3. 器皿　转化实验使用的器皿如玻璃器皿、微量吸头及 Eppendorf 管等，应彻底清洗和高压消毒，表面去污剂及其他化学试剂污染往往大幅度降低转化率。

4. 感受态细菌 4℃ 保存 12～24h，可提高转化率。

（朱法良　梁晓红）

实验八　重组子的筛选与鉴定

Exp8. Screening and Identification of Recombinants

如何从转化细菌菌落中筛选含有阳性重组子的菌落并鉴定重组子的正确性，是基因工程的重要工作。不同的克隆载体及相应的宿主系统，其重组子的筛选、鉴定方法不尽相同，概括起来有以下几类。

一、重组子的筛选法

重组子转化宿主细胞后，载体上的筛选标志基因会导致细菌某些表型的改变，若在琼脂平板中添加相应的筛选物质，可以直接筛选鉴别含阳性重组子的菌落，常是筛选阳性重组子的第一步。

1. 抗生素平板筛选　为阳性重组子的初步筛选方法。对带有抗生素抗性基因的载体，如果外源 DNA 片段插入载体的位点在抗性基因之外，抗性基因仍能表达，这样重组子转化细胞后，能在含有相应药物的琼脂平板上生长并形成菌落的，是含有阳性重组子的宿主菌。

2. 插入失活双抗生素对照筛选　有的载体含有双抗性基因，外源 DNA 插入后使其中一个抗性基因失活，则可用两个分别含不同药物的平板对照筛选出阳性重组子。

3. 插入表达筛选　有些载体的筛选标志基因前面存在一段负控制序列，当插入失活该负调控序列时，筛选标志基因才能表达，为插入表达筛选。

4. β-半乳糖苷酶筛选　对携带 *lacZ* 基因的载体，可以利用与宿主菌的 α-肽互补作用进行筛选。*lacZ* 基因编码 α-肽，可与宿主菌形成 α-肽互补作用后，编码有活性的 β-半乳糖苷酶，通过 IPTG 诱导，分解底物 X-gal 形成蓝色噬菌斑。若在 *lacZ* 基因的多克隆位点上插入外源基因，可使 *lacZ* 基因失活，从而破坏重组子与宿主菌之间的 α-肽互补作用，在含有 X-gal 和 IPTG 的平板中生长时形成无色噬菌斑或菌落，而载体自身环化后转化的细菌为蓝色噬菌斑或菌落。此法筛选出阳性重组子的几率比较高。方法如下：

(1) 制作含 X-gal 的琼脂平板：含 1.5% 琼脂的 LB 固体培养基融化后，冷却至 40℃，加入氨苄青霉素使终浓度为 50μg/ml，混匀后铺琼脂平板，待凝固后，再在平板表面均匀涂布 20μl 50mg/ml 的 X-gal 和 100μl 100mmol/L 的 IPTG，使其终浓度分别为 40μg/ml、0.5mmol/L。

(2) 取适量的转化菌，均匀涂布上述平板，37℃ 生长 24h，观察形成的菌落。

5. 快速裂解菌落法　从抗性平板中直接挑取菌落裂解后，不需要内切酶消化，直接进行凝胶电泳，与载体 DNA 比较迁移率，初步判断是否有插入片段的存在，适用于插入片段

较大重组子的初步筛选。方法如下：

(1) 用消毒牙签从菌落上挑取少许细菌接种到另一个 LB 琼脂板，再将该菌落中剩下的细菌全部取出至 96 孔培养板中，加入裂解液，用牙签打散菌落。

(2) 盖上 96 孔板的盖，用胶布封严，68℃ 水浴保温 45min。

(3) 每孔加入 2.5μl 25%蔗糖，对照组为载体 DNA，上样于琼脂糖凝胶进行电泳。

(4) 电泳完毕后，凝胶于含有 0.5μg/ml EB 的溶液中染色 45min，紫外线灯下观察，如有比载体 DNA 迁移慢的样品，则根据编号，从储存琼脂板上，挑取相应的菌落，小量培养提取 DNA，进行内切酶消化鉴定。

二、重组子的鉴定

一般情况下，经过初步筛选的重组子，还需根据重组子的结构特征进行酶切、杂交、PCR 扩增等步骤，才能鉴定阳性重组子。

1. 内切酶图谱鉴定　对于初步筛选出的阳性重组子，应用相应的内切酶切割释放出目的基因片段，然后电泳检测插入片段和载体的大小。对于可能存在双向插入的重组子，还可以用内切酶消化后鉴定方向。

2. Southern 印迹杂交　为进一步确定 DNA 插入片段的正确性，在内切酶消化重组子、凝胶电泳后，通过 Southern 印迹杂交，鉴定重组子中的插入片段是否是所需的靶基因片段。具体方法参见核酸杂交技术。

3. PCR 筛选重组子　利用载体的外源 DNA 插入位点两侧设计引物，对初步筛选的阳性重组子进行 PCR 分析，不但可迅速扩增插入片段，而且可以鉴定插入片段的序列正确性，判断连接的方向。

4. 菌落原位杂交　菌落或噬菌斑原位杂交技术是至今仍通用的重组子筛选技术，它是先将转化菌直接铺在硝酸纤维素膜上，用放射性核素或地高辛标记的特异 DNA 或 RNA 探针进行分子杂交，然后挑选出阳性克隆菌落的方法。本方法能进行大规模操作，一次可筛选 $5\times10^5 \sim 5\times10^6$ 个菌落或噬菌斑，对于从基因文库中挑选目的重组子，是一项首选的方法。

(1) 抗性平板的制作　含 1.5%琼脂的 LB 固体培养基融化后，冷却至 40℃，加入氨苄西林使终浓度为 50μg/ml，混匀后铺琼脂平板。

(2) 取两块含有 Amp 的 LB 琼脂板，用剪刀剪取适当大小(直径小于平皿内径)的硝酸纤维素膜，光面向上铺于其中一块含有 Amp 的 LB 琼脂板上，做好标记，待硝酸纤维素膜湿润后，备用。

(3) 用消毒牙签从抗性平板上挑取单菌落，点至硝酸纤维素膜上，然后再用此牙签将该菌落接种到另一个含有 Amp 的 LB 琼脂板的相应位置处。

(4) 换消毒牙签，依此法再接种其他的单菌落。

(5) 将平板放入 37℃ 生长 24h，取下硝酸纤维素膜，并保存好另一平板。将硝酸纤维素膜经过变性液变性、中和液中和后，80℃ 真空烤干膜 2h，与相应的 DNA 或 RNA 探针杂交，根据杂交结果，于另一平板中挑取相应的阳性菌落，活化后抽提质粒或冻存菌种。

以上介绍了筛选鉴定阳性重组子较为常用的方法，在具体的实验操作中，可根据所用的

质粒类型、插入片段的大小、宿主菌的特点等实际情况选用合适的一种或几种方法。

（刘素侠　张　秋）

实验九　DNA 探针标记及核酸杂交技术

Exp9. DNA Labeling Technology and Nucleic Acid Hybridigation

一、DNA 探针标记与斑点杂交

（一）**DNA** 探针标记

【原理】

随机引物法是一种较理想的核酸探针标记方法，其基本原理是，按照碱基互补配对原则，随机引物(6 聚核苷酸引物)与待标记的 DNA 片段上的不同区段特异性结合，在 Klenow 大片段的催化下，即从引物 3′-OH 端开始合成互补 DNA 链，反应中若加入 Dig(digoxigenin，地高辛)标记的脱氧尿嘧啶三磷酸(Dig-dUTP)，其就可随机掺入到新合成的 DNA 链中，即可获得 Dig 标记的探针。

本法的优点是标记速度快(1h)；标记量宽(10ng～3μg)；检测反应灵敏度高；线性 DNA 标记效率高；没有半衰期，安全性好；适用范围广(各种杂交)。

【材料】

待标记 HBV DNA。

6 聚核苷酸引物。

dNTP 标记混合液：10 × dNTP 标记混合液(含有 dATP、dCTP、dGTP、dTTP；Dig-dUTP，pH7.5)。

Klenow 酶：2u/μl。

TE 液：10mmol/L Tris-HCl，1mmol/L EDTA。

【方法】

以 HBV DNA-Dig 探针的标记为例，方法如下：

变性 HBV DNA	5μl
6 聚核苷酸引物	2μl
10×dNTP(含 Dig-dUTP)	2μl
Klenow 酶	1μl
灭菌双蒸水	10μl
总体积：	20μl

1. 取一 1.5ml Eppendorf 管，加入 HBV DNA 2μg(10μl)。

2. 沸水中煮沸 10min，迅速放入冰水浴中冷却 10min。

3. 另取一 1.5ml Eppendorf 管，分别加入：各种试剂(见上表)。

4. 轻度混匀，12 000r/min 离心 10s，37℃ 水浴过夜。

5. 加入 0.2mol/L EDTA 2μl 终止反应，再加入 2.5μl 的 3mol/L NaAc，75μl 的冷无水乙醇，混匀，-20℃过夜。

6. 14 000r/min 4℃离心 20min，用 1ml 预冷的 70%乙醇溶液洗一次，14 000r/min 4℃离心 20min。

7. 弃上清，室温干燥后以 50μl TE 重溶，置于-20℃保存待用。

（王晓燕　张艳）

（二）斑点杂交(dot blot hybridization)

【原理】

将待测 DNA 样品经变性后，点样于硝酸纤维素膜上，用 Dig 标记探针与其进行杂交。若待测样品中有与探针同源的 DNA 存在，标记的探针即与待检 DNA 按碱基配对原则结合，再与碱性磷酸酶标记的 Dig 抗体结合，在碱性磷酸酶底物存在的情况下，可显现出褐色斑点，为阳性结果。

Ⅰ. 探针灵敏度的检测

【材料】

1μg HBV DNA 样品、Dig-HBV 探针。

20×SSC：NaCl 175.3g；枸橼酸钠 88.2g；双蒸水 800ml，用 NaOH 调 pH 至 7.0，加水至 1000ml，高压灭菌。

预杂交液：5×SSC；5%封阻剂；50%甲酰胺等。

洗膜 I(5×贮存液)：10×SSC；0.5%SDS(使用时稀释)。

洗膜 II(5×贮存液)：1×SSC；1%SDS(使用时稀释)。

缓冲液 A(10×贮存液)：100mmol/L Tris-HCl；1.5mol/L NaCl 稀释使用。

缓冲液 B：0.5%封阻剂溶液，-20°C 贮存。

缓冲液 C：100mmol/L Tris-HCl；100mmol/L NaCl；50mmol/L $MgCl_2$，pH9.5。

抗-Dig-AP 结合物；显色剂Ⅰ：NBT；显色剂Ⅱ：X-磷酸盐。

硝酸纤维素膜、杂交袋等。

【方法】

1. 硝酸纤维素膜的预处理　将适当大小的硝酸纤维素膜用蒸馏水浸透，然后浸于 20ml 10×SSC 溶液中 5min 以上，再将硝酸纤维素膜铺于 96 孔点样器上。

2. 样品与点样　用灭菌双蒸水将 1μg HBV DNA 样品稀释成系列梯度，使每个梯度 HBV DNA 的浓度依次为 10ng、1ng、100pg、10pg、1pg、0.1pg、0.01pg/10μl。每梯度取 10μl，100℃煮沸 5～10min，立即放入冰水浴中冷却 5min。通过 96 孔点样器吸印到硝酸纤维素膜上，真空泵负压抽干。每梯度点两孔，如下图：

3. 烤膜　取下硝酸纤维素膜放在滤纸上，放入 80℃温箱，烤干 2h。

4. 预杂交　取出膜放于室温，将膜于 10×SSC 中浸湿后，放入塑料袋中预杂交，加入 2ml 预杂交液，赶出气泡封口，42℃水浴摇床 2～3h(可手动翻动几次)。

1	2	3	4	5	6	7
○	○	○	○	○	○	○
○	○	○	○	○	○	○

5. 杂交　倒净杂交袋中的溶液，加 1ml 新杂交液，再加入已加热煮沸(100℃ 10min)并冷却变性好的(冰浴 5min)探针 10μl，小心赶出气泡封口，混匀袋内溶液，放 42℃水浴摇床摇动过夜(>20h)。

6. 洗膜　先按比例稀释好洗膜Ⅰ和洗膜Ⅱ。用适量 1×洗膜Ⅰ溶液室温下振摇洗 2 次，每次 5min。再用适量 1×洗膜Ⅱ溶液于 50～55℃下振摇洗 2 次，每次 15min。

7. 显色反应

(1) 用蒸馏水 10 倍稀释缓冲液 A，膜在适量 1×缓冲液 A 中短暂洗涤 1min。

(2) 膜在适量缓冲液 B 中，室温摇动 30min。

(3) 再用适量 1×缓冲液 A 短暂洗膜。

(4) 用 10ml 1×缓冲液 A 稀释 100μl 酶结合物(1∶100 稀释)，室温摇动 30min。

(5) 用适量 1×缓冲液 A 摇动洗膜(以去除未结合的抗-Dig-酶结合物)室温 15min，洗 2 次。

(6) 膜在适量缓冲液 C 中平衡 2min 后，倒净溶液。

(7) 在 10ml 缓冲液 C 中，加入 45μl NBT 和 35μl X-磷酸盐溶液，混匀后在黑暗下显色，阳性对照应在 30min 内显色，显色时不要摇动膜。

(8) 在显色适当时(一般 5h)，取出膜用 TE 溶液终止显色反应。延长显色反应的时间可提高一定的灵敏度。

8. 结果判定　一般可直接检测出 0.1pg 的 DNA 量。若低于 0.1pg 则可能由于 DNA 定量不准或试剂盒标记效率不高。

【注意事项】

1. 标记好的探针除做灵敏度鉴定外，同时应做含有阴性、阳性对照的探针特异性检测。

2. 可直接将标记好的探针系列稀释后点样于硝酸纤维素膜上，然后直接进行显色反应，这样可省去前面预杂交和杂交的步骤，但不宜定量。

Ⅱ. 检测血清标本中的 HBV DNA

【材料】

蛋白酶 K 缓冲液：0.4mol/L NaCl；100mmol/L Tris；1mmol/LEDTA；4%SDS。

蛋白酶 K：20mg/ml。

饱和酚、氯仿/异戊醇(24∶1)。

【方法】

1. 取 1.5ml Eppendorf 管，加入 10μl 蛋白酶 K、50μl 蛋白酶 K 缓冲液、50μl 患者血清，37℃水浴 2h。

2. 加 110μl 酚：氯仿/异戊醇(1∶1)混合液于上述离心管中，混匀 5min，12 000r/min 离心 10min。

3. 在新离心管中加入 3μl 15mol/L NaCl,吸取离心好的上清 60μl,加 150μl 冷无水乙醇,混匀,-20℃过夜或 2h,12 000r/min 离心 10min,沉淀核酸。

5. 弃上清,真空干燥(约 15min)。每管加入 50μl TE 液溶解核酸。

6. 点样:将已提取好的血清核酸和阳性、阴性对照各 50μl 于 100℃煮沸 5~10min,立即放入冰水浴中冷却 5min。短暂离心使管壁的溶液离心到管底,再加入 50μl 的 10×SSC 混匀后,取 100μl 通过 96 孔点样器吸印到硝酸纤维素膜上,真空泵负压抽干。

杂交步骤同探针灵敏度的检测。

【注意事项】

1. 在预杂交和杂交过程中溶液应处于流动状态,替换杂交时膜不能干燥。

2. 杂交后,含探针的杂交液应回收,-20℃保存,下次变性后可以重复使用。

3. 膜可以立即用于显色,或放在干燥的环境中待以后显色用。

(王晓燕　高立芬)

二、Southern 印迹杂交

【原理】

Southern 印迹杂交(southern blot hybridization)是将 DNA 限制性内切酶酶切片段在琼脂糖凝胶中电泳分离后,在高盐条件下通过虹吸作用将变性的 DNA 片段转印到硝酸纤维素膜上,再利用特异性探针进行杂交。Southern 印迹杂交是对基因或 DNA 片段进行定性、定量、突变定位及定长的常用方法。在遗传病、DNA 图谱分析、基因突变、RFLP 及 PCR 产物分析方面都有重要应用。

转移 DNA 除了应用虹吸作用外,还可用电转法或真空转移法。

【材料】

0.25mol/L HCl

变性液:0.5mol/L NaOH,1.5mol/L NaCl。

中和液:1mol/L Tris-HCl(PH8.0),3mol/L NaCl。

硝酸纤维素膜、普通滤纸、Whatman 3MM 滤纸、吸水纸、玻璃板,500g 砝码(或重物),其他试剂同斑点杂交。

【方法】

1. 经酶切的 DNA 标本在琼脂糖凝胶中电泳后,EB 染色,在紫外光下摄影,电泳时,应有一孔加入 *Hin*dⅢ酶切的 λDNA 作为 DNA 相对分子质量的对照。

2. 将凝胶置一大平皿中,切去无用区域,并在胶的一角作好标记。

3. 酸处理:在转移大分子量 DNA(>1kb 时)时,常需用 HCl 处理以打碎 DNA,加入 0.25mol/L HCl 使液面没过胶面,置室温处理 10~20min,轻轻摇动数次,倒去液体,用蒸馏水洗 3 次。

4. 变性:加入变性液置室温 30min,使双链 DNA 间的氢键打开,轻轻摇动数次,倒去液体,用蒸馏水洗 3 次。

5. 中和：加入中和液置室温 30min，重复一次。

6. 用 Whatman 3MM 滤纸覆盖在玻璃板上，将玻璃板置一个小瓷盘上，加入 10×SSC，用玻璃棒赶去滤纸与玻璃之间的气泡。

7. 将凝胶轻轻倒置于滤纸上，用玻棒赶尽气泡。

8. 将稍小于凝胶大小的硝酸纤维素膜在 2×SSC 内浸湿后(5min 以上)，覆盖于凝胶之上，小心除去滤膜与凝胶之间的气泡，在凝胶周围包上一层保鲜膜，防止液体流动形成短路，然后再覆盖同样大小 Whatman 滤纸 8 张。

9. 将吸水纸裁成与凝胶相当大小，置于滤纸之上(5～8cm 高)。再加一块玻璃板，玻璃板上放一 500g 重的砝码(或其他重物)，液体可从"凝胶→硝酸纤维素膜→3MM 滤纸"方向流动，从而使凝胶中的 DNA 片段转移至硝酸纤维素膜上。

10. 转移约需 12～24h，转移期间，当吸水纸潮湿时应及时替换。

11. 移去吸水纸，取下硝酸纤维素膜，置 2×SSC 中漂洗后取出，置 2 层滤纸中，80℃ 真空烘干 2h。置于燥缸内保存，待杂交。

12. 其他步骤同斑点杂交。

【注意事项】

1. 在处理凝胶时动作要轻，以防凝胶破碎。

2. 重物可使滤纸片、吸水纸之间有较好的接触，发挥毛细虹吸作用。但注意不要太重，这样会影响 DNA 从胶中向上转移。

3. 转移时间取决于 DNA 片段的大小。若片段＜1kb 转移 1～2h 即可，若＞15kb 需转移 15h 以上。

(王晓燕　张蘋)

三、Northern 印迹杂交

【原理】

Northern blot 印迹杂交(Northern blot hybridization)与 Southern blot 杂交基本相同。所不同之处在于，RNA 不如 DNA 稳定，易被 RNA 酶降解，而 RNA 酶不仅广泛存在，而且加热至 100℃ 也不能使其灭活，所以在操作 RNA 时，所用的一切器皿和溶液都必须经过严格的消除 RNA 酶的处理。另外，RNA 的电泳必须在含甲醛或戊二醛变性凝胶中进行。Northern blot 杂交主要应用于检测癌基因或其他肿瘤相关基因的过度表达，也可检测抑癌基因的低表达或不表达。

【材料】

10×MOPS 电泳缓冲液：0.2mol/L MOPS(3-[*N*-吗啉]丙磺酸)(pH7.0)，0.05mol/L 乙酸钠，0.01mol/L EDTA(pH8.0)。

载样缓冲液(用 DEPC 预处理)：50% 甘油，1mmol/L EDTA(pH8.0)，0.25% 溴酚蓝，0.25%二甲苯青。

普通琼脂糖，37% 甲醛，甲酰胺。

电泳装置,相对分子质量标准参照物,其他试剂同斑点杂交。

【方法】

1. 配制 1%琼脂糖变性胶:

10×MOPS 电泳缓冲液	10ml
0.1% DEPC 水	90ml
普通琼脂糖	1g

将胶加热至完全熔化后，在室温放置冷却至 60℃，加入 37%甲醛 5.4ml,混匀。然后灌注电泳胶，电泳槽必须用去污剂洗干净，蒸馏水冲洗，乙醇干燥，然后用 3%双氧水处理 10min，最后用 DEPC 处理的蒸馏水彻底冲洗。

2. 在一个 1.5ml Eppendorf 管中加入下列试剂并混匀:

RNA	4.5μl
5×MOPS 电泳缓冲液	2.0μl
37%甲醛	3.5μl
甲酰胺	10.0μl

置 65℃保温 15min，速置冰浴中，12 000r/m,离心 30s,加入 2μl 载样缓冲液,混匀。

4. 上样前预电泳 5min,将 RNA 样品上样于加样孔中，同时在另一样品孔中加入相对分子质量标准参照物，恒压 3～4V/cm 电泳。每 1～2h 将阴阳极电泳液混合一次。

5. 电泳结束后，切下部分相对分子质量标准参照物带，溴化乙锭染色，紫外线下照相或凝胶成像仪下拍照,记录结果。

6. 上述电泳后的凝胶一般不需进行处理即可直接进行印迹。含有甲醛的凝胶可用 DEPC 预处理过的水漂洗以去除所含的甲醛。如果凝胶较浓(>1%)，较厚(>0.5cm)或待测 RNA 片段较大(>2.5kb)，可预先将凝胶置 0.05mol/L NaOH 溶液中浸 20min，再用 DEPC 处理的水漂洗，最后用 20×SSC 浸 45min。转移方法同 Southern 印迹。

【注意事项】

所用的器皿及试剂均必须经过 DEPC 浸泡,再高压后高温烘烤。

（王晓燕　张蘋）

实验十　从组织中提取人基因组 DNA

Exp10. Extraction of Human Genome DNA From Tissue

【原理】

在进行真核基因工程、Southern 杂交或 PCR 时,需从组织或细胞中获得完整的基因组 DNA,因而真核生物的组织或细胞(包括培养细胞)常成为制备 DNA 的主要材料。

提取基因组 DNA 的一般原理是取冰冻或液氮冻结的组织,经剪碎、匀浆器研碎或用工具砸碎以后,经 SDS 和蛋白酶 K 作用使细胞充分裂解释放出 DNA,然后用饱和酚和氯仿/异戊醇去除蛋白质,用 RNA 酶去除 RNA,就可以产生 100~200kb 左右的基因组 DNA 片段,经适当剪切后,可适用于以 λ 噬菌体作为载体的基因组文库的构建。

【试剂】

PBS。

匀浆缓冲液:0.25mol/L 蔗糖;25mmol/L Tris-HCl(pH7.5);25mmol/L NaCl;25mmol/L $MgCl_2$。

细胞裂解缓冲液(TNE):10mmol/L Tirs-HCl(pH7.4);10mmol/L NaCl;10mmol/L EDTA。

10%SDS;Proteinase K(10mg/ml)。

饱和酚(pH8.0),氯仿/异戊醇(24:1)。

3mol/L NaAc,无水乙醇等。

【方法】

1. 取小鼠新鲜肝组织,在电子天平秤上称取 0.1g,放置于青霉素小瓶中备用或 -20℃ 存放。

2. 用一锐利剪刀将组织剪成尽可能小的块(处理过程中要保持组织的潮湿)。加入 1ml PBS 混匀。

3. 将剪好的组织移至 1.5ml Eppendorf 管内,5000r/min,离心,2min。

4. 弃上清,用冷匀浆缓冲液 1ml 悬浮沉淀,转移至匀浆器内,于冰上研磨 15~20 次(不能超过 30 次)。

5. 再将研磨好的组织移至新的 1.5ml Eppendorf 管内,5000r/min,离心,2min。

6. 弃上清,用 1ml PBS 悬浮沉淀,5000r/min,离心 2min。

7. 弃上清,用 1ml 细胞裂解缓冲液悬浮沉淀,加入 50μl 10%SDS 使终浓度为 0.5%,轻轻混匀后,加入 20μl 10mg/ml 的 Proteinase K 至终浓度为 200μg/ml,充分混匀。

8. 37℃ 水浴过夜后,分至两个 1.5ml Eppendorf 管内

9. 加入等体积酚/氯仿/异戊醇(25:24:1)混合，抽提 10min。

10. 12 000r/min，离心 15min，吸出水相至另一新 1.5ml Eppendorf 管中。如果中间变性蛋白层较厚，再加入 1/4 体积的细胞裂解缓冲液至含酚和中间层的管中，如上步骤再提取一次，合并两次的水相。

11. 重复第 9、10 步。

12. 取上清液至另一新 1.5ml Eppendorf 管中，加入 1/10 体积 3mol/L NaAc 和 2 倍体积无水乙醇，混匀，可见 DNA 呈絮状析出，此时可用玻璃棒将 DNA 绕出溶于 TE 中，剩余液体可放置 -20℃ 2h 或过夜。

13. 12 000r/min，离心 10min 使 DNA 沉淀，沉淀用 70% 乙醇漂洗一次，所得 DNA 样品溶于 50μl TE 中。

【注意事项】

1. 组织块取材要新鲜，应尽量剪碎。

2. 所用匀浆器须配套适中，过大过小均不利于研磨组织块。

3. SDS 在加样前要求溶解，可将其置于 37℃ 水浴中促溶。

4. 尽可能使用新鲜配制的蛋白酶 K，使用前需进行预试验确定其活性。

5. 加入等体积酚/氯仿/异戊醇(25:24:1)混合抽提时应颠倒混匀，要轻柔，避免剧烈震荡。

6. 苯酚腐蚀性很强，可引起严重的烧伤，注意防护。

(王晓燕　张　艳)

实验十一　基因转染哺乳动物细胞

Exp11. Gene Transfection into Mammalian Cells

转染(transfection)是指将外源DNA或RNA导入细胞的过程。哺乳动物细胞常规的转染方法有两种:短暂转染和稳定转染。前者又称为过渡性转染、瞬时转染,是将DNA送入活细胞内,但DNA不会嵌入细胞的染色体中,可在2~3天内通过收集被转染的细胞进行基因表达分析。后者又称为永久性转染,是将目的基因传送入活细胞,经过一定时间选择标记的筛选,使目的基因整合入细胞染色体中或者成为基因附体存在于细胞内,从而得到稳定表达目的基因的细胞系。

目前,基因转染技术常被用于基因表达的分析研究上,常用的转染方法有:①磷酸钙(calcium-phosphate)转染法,②DEAE-葡聚糖转染法(DEAE-dextran),③脂质体(liposome)介导的转染法,④电穿孔法(electroporation),⑤显微注射法(microinjection),⑥病毒载体介导的转染法,⑦基因枪(gene gun)转染法,⑧受体介导的基因转染法。现将常用的几种转染方法介绍如下。

一、电穿孔法

【原理】

细胞暴露于高电压时,细胞膜可被电脉冲瞬间穿孔,从而使存在于细胞悬液中的外源DNA进入细胞。

【材料】

待转染的哺乳动物细胞。

不含及含有选择剂的完全培养基。

电穿孔缓冲液(冰冷)。

纯化的DNA。

Beckman JS-4.2转子或相当转子。

电穿孔用的电极池及电源。

【方法】

1. 在完全培养液中培养待转染细胞至对数生长期,4℃,640r/min离心5min收集细胞。

2. 将细胞沉淀用其半量体积的预冷电穿孔缓冲液重悬洗涤,4℃,640r/min离心5min。

3. 对于稳定转染,将细胞以1×10^7/ml的密度重悬于0℃的电穿孔缓冲液中;对于短暂表达,可用更高密度的细胞(高达8×10^7/ml)

4. 在冰上放置所需数量的电穿孔用的电极池，每池中移入 0.5ml 细胞悬液。

5. 将 DNA 加入冰上各电极池的细胞悬液中。握住电极池两侧的窗口，晃动其底部以混匀 DNA/细胞悬液，在冰上放置 5min。

6. 将电极池放入电穿孔装置的架上（室温），用所设定的电压及电容值电击一次或多次。

7. 将电极池至于冰上 10min。

8. 用非选择性完全培养液 20 倍稀释被转染的细胞，并用该液洗电极池以移出所有的被转染细胞。

9. 对于稳定转染，在非选择性培养液中培养细胞 48h（大约两代），然后转入含有抗生素的培养液中；对于短暂转染，培养细胞 50～60h 收集细胞进行短暂表达分析。

【注意事项】

1. 电穿孔法受 DNA 浓度影响较小，一般 DNA 的量在 10～40μg/10^7 细胞的范围内效果好，并且在 DNA 用量与摄入量之间有好的线性关系。

2. 电穿孔法中可变的参数是电泳冲的强度和长度。操作中最好能找到杀死 20%～60%细胞的电脉冲（一般是 1.5kV，25μF）。如果过多的细胞发生死亡，可通过调低电容来降低脉冲的长度。

二、脂质体介导的转染

【原理】

脂质体（lipofectin）是一种人工制备的类脂质小球体，有一个或多个酷似细胞膜的类脂质分子包裹着水相介质组成。脂质体主要组成成分是天然形成的磷酸类脂。构成双分子层的类脂其亲水性的首基部分形成膜的内外表面，而亲脂性的尾端部分位于膜的中间。膜壁的厚度约为 5～7nm，而囊的直径一般在 25～500nm 之间。脂质体的这种结构使其能够携带各种亲水的、疏水的及两亲的物质，它们被包入脂质体内部水相、插入类脂双分子层或吸附连接在脂质体的表面。按荷电性，脂质体可分为中性脂质体、负电性脂质体、阳离子脂质体。

阳离子脂质体，又称阳性脂质体、正电荷脂质体，是目前研究较为详尽的一种转染脂质体。大多数阳离子脂质体是由一种中性磷脂和一种或多种阳性成分组成，其中，中性磷脂起到稳定双层膜和降低阳性成分毒性的作用，同时提高阳性脂质的细胞渗透功能。阳性成分是具有不同化学结构的两性分子，多为双链季铵盐型表面活性剂，为整个脂质体提供正电荷。作为本身带有正电荷的脂质囊泡，可作为荷负电物质的传递载体，特别是适用于蛋白质、多肽和寡核苷酸类物质（DNA 或 RNA）。阳离子脂质成分与带负电荷的 DNA 相互作用，进而将 DNA 缩合成致密结构，形成脂质-DNA 复合物。阳离子脂质体-DNA 复合物与胞质膜直接融合，将 DNA 释放到胞质中。

脂质体介导的基因转染有以下特点：①脂质体与基因的复合过程比较容易；②脂质体与细胞膜融合将目的基因导入细胞后，脂质体即被降解，无毒，无免疫原性；③DNA 或 RNA 可得到保护，不至于灭活或被核酸酶降解；④转染过程方便易行，重现性好。

【材料】

指数期生长的 HepG2 细胞。

纯化的质粒 pcDNA-GFP。

DMEM 培养基，不含血清。

含 10％胎牛血清的完全培养液（DMEM/10％胎牛血清）。

脂质体悬液（LipofectAMINE™2000 Reagent/LF2000）。

24 孔细胞培养板。

5％CO_2，37℃的加湿培养箱。

聚苯乙烯管（Falcon 或 Corning）。

【方法】

1. 转染前一天，按 $1\times10^5\sim3\times10^5$ 细胞/孔的量在 24 孔细胞培养板中接种指数期生长的细胞，在 5％CO_2，37℃的加湿培养箱中培养过夜，直至细胞 90％～95％汇片。

2. 在 Eppendorf 管中配制 DNA/脂质体混合物：对于每孔细胞，将 0.8～1.0μg 质粒以无血清培养基稀释，终体积为 50μl；以相同无血清培养基稀释脂质体悬液至 50μl，每孔细胞需脂质体 1～3μl；将两管稀释液在室温 5min 之内混匀至一管中，混匀时应逐批进行，不断吹打使二者充分融合，室温放置 20min。

3. 取前一日接种培养板，移去旧的培养液，将 DNA/LF 2000 混合物（100μl）加至细胞孔的中央，并同时加入 0.5ml 无血清培养基，前后晃动培养板数次，使 DNA/LF 2000 混合物均匀分布至细胞表面。5％CO_2，37℃的加湿培养箱中培养 4～6h。

4. 4～6h 后，更换含 10％胎牛血清的相应培养基，继续培养 24～48h。

5. 若原细胞培养为含抗生素的培养基，更换后继续培养 24h。

6. 若为短暂转染，转染后 72h 即可收集细胞用于表达分析；若为稳定转染，则需在恢复培养条件 24h 后，按 1:10 的比例传代细胞，并更换选择培养基，直至得到阳性细胞克隆。

【注意事项】

1. 脂质体对细胞有一定的毒性，因此转染时对细胞的浓度要求较高，而且不同的细胞敏感性不同，可通过预实验优化转染细胞数量。

2. 脂质体悬液被稀释好之后，最好能在 5min 之内与待转染 DNA 混合，不宜超过 30min。否则脂质体活力会大大降低。

3. 为使待转染 DNA 与脂质体充分混匀，应逐步进行，如每次取 10μl。

4. 为提高脂质体转染效率，转染时细胞应在无血清、无抗生素的条件下培养，警惕细胞污染的发生。

（马春红　梁晓红）

实验十二　蛋白质印迹技术

Exp12. Western Blot Hybridization

1975 年 Southern 创建了 DNA 印迹技术(Southern blot),在此基础上,1979 年和 1981 年 Towbin 和 Burnette 分别发表文章首次建立了检测蛋白质的免疫印迹技术(Western blot)。Western blot 具有与固相免疫测定相似的灵敏度和特异性,但却没有放射免疫测定的放射性污染,也不需要像免疫沉淀一样标记靶蛋白。Western blot 可检测出 1～5ng 的蛋白质。是从混合物中定向、定量检测特异蛋白的有效手段。

【原理】

Western blot 由 SDS-PAGE 电泳、蛋白质转印和固相免疫测定三项技术结合而成:首先蛋白质经 SDS-PAGE 电泳而据相对分子质量大小分离成不同条带;分离蛋白质通过电转移/直接印迹方式原位转移至固相支持物上,并保持其生物学活性不变;在固相支持物上让相应抗体与被分离蛋白质作用,根据抗原抗体结合的特异性检测靶蛋白。

(一) SDS-PAGE 电泳

SDS-PAGE 电泳具有三种物理效应:分子筛效应、浓缩效应和电荷效应,可使蛋白质按相对分子质量大小精确分离。在电泳前,蛋白质样品应先经 100℃ 煮沸 3min,使之变性,防止多聚体形成。

【材料】

30%聚丙烯酰胺凝胶贮备液:温热去离子水配制成含 29%(W/V)丙烯酰胺(Arc)和 1%(W/V)亚甲双叉丙烯酰胺(Bis)的凝胶贮存液(pH≤7.0),室温贮存于棕色瓶中,有毒性。

TEMED(*N*,*N*,*N*′,*N*′-四甲基乙烯二胺):催化作用,可加速凝胶聚合。

10%过硫酸铵:贮存于 4℃,在一周内使用,可加速 Bis 和 Arc 聚合。

Tris-Aly 电泳缓冲液:25mmol/LTris,250mmol/L 甘氨酸(pH8.3),0.1%SDS。

上样缓冲液:50mmol/L Tris(pH6.8), 100mmol/L 二硫苏糖醇(DTT),2%SDS, 0.1%溴酚蓝、10%甘油。

1.5mol/LTris 缓冲液(pH8.8),1.0mol/L Tris 缓冲液(pH6.8),10%SDS。

【方法】

1. 安装垂直电泳槽的两块玻璃板。

2. 根据分离蛋白质相对分子质量大小,选择适当的聚丙烯酰胺凝胶浓度(表实验 12-1),并据表实验 12-2 数值在三角烧瓶中按所需凝胶浓度配制一定体积的分离胶溶

表实验 12-1　丙烯酰胺浓度与蛋白质分离范围

丙烯酰胺浓度/%	线性分离范围/($\times 10^3$)
15	12～43
10	16～68
7.5	36～94
5.0	57～212

液，一旦加入 TEMED,应立即混匀并灌胶。

表实验 12-2 配制 15ml 各种浓度丙烯酰胺所需各成分的体积(ml)

凝胶浓度	溶液成分					
	水	30%Arc	1.5mol/L Tris (pH8.8)	10%SDS	10 %过硫酸铵	TEMED
6%	7.9	3.0	3.8	0.15	0.175	0.056
8%	6.9	4.0	2.8	0.15	0.175	0.042
10%	5.9	5.0	2.8	0.15	0.175	0.028
12%	4.9	6.0	3.8	0.15	0.175	0.028
15%	3.4	7.0	3.8	0.15	0.175	0.028

3. 迅速在玻璃间灌胶,留出成层胶空间(梳子齿长 + 1cm),小心在分离胶上覆盖一层无离子水或 0.1%SDS,室温垂直放置 2～3h。

4. 完全聚合后,倒出覆层液体,尽可能排去液体，用吸水纸吸去残留液体(不可触破胶面)。

5. 制备成层胶:按表实验 12-3 配制成层胶，一旦加入 TEMED，应立即混匀并灌胶。

表实验 12-3 配制 6ml 成层胶所需各成分的体积(ml)

水	30%丙烯酰胺	1mol/LTris(pH6.8)	10%SDS	10%过硫酸铵	TEMED
4.1	1.0	0.75	0.06	0.06	0.012

6. 在已聚合的分离胶上直接灌注成层胶，立即插入样品梳(小心避免气泡产生),垂直室温放置。

7. 聚合时,将样品置于上样缓冲液中,100℃ 3min 使蛋白质变性。

8. 聚合后，小心拔出梳子,将电泳槽中灌满 Tris-Aly 电泳缓冲液,并用电泳缓冲液冲洗样品孔。

9. 将变性后的样品小心加至样品孔底部。

10. 接电源(正极接下槽),小电流(15mA)使染料进入分离胶后,再加大电流至 20mA,继续电泳至溴酚蓝到达分离胶底部，关闭电源。

11. 卸下玻璃板，小心取出凝胶。

(二) 蛋白质转印

通过电转移或直接印迹法将蛋白质条带原位转移至固相支持物上。常用的蛋白质转印方法是电转移，即通过电泳将 SDS-PAGE 凝胶上的蛋白原位、精确地转移至固相支持物上。Western blot 中作为蛋白质的固相支持物有很多，最常用的是硝酸纤维素(NC)膜,它可与蛋白质非共价结合。

【材料】

转移液:25mmol/L Tirs(pH8.0)，192mmol/L Aly,20%甲醇。

0.5%氨基黑水溶液

10%乙酸。

封闭液:含 3%牛血白蛋白，0.5%Tween-20 的 PBS

【方法】

1. 电泳后将凝胶放至转移液中平衡。

2. NC 膜处理　双蒸水中漂洗 2h，转到转移液中 4h。Whatman 滤纸同 NC 膜处理。

3. 转膜　在海绵夹中依次放置两层 Whatman 滤纸、凝胶、NC 膜、两层滤纸，将膜朝向正极，胶向阴极置于转移槽中（槽中注满转移液），100mA 转移 15～18h。

4. 转膜后，切下 NC 膜边缘的对照蛋白质，在 0.5% 氨基黑水溶液中染色 3s，10% 乙酸脱色，若 NC 膜上有蛋白质带，证明转膜成功。

5. 转移后的 NC 膜空气中干燥 0.5h 后，置于封闭液中 4℃ 封闭 7～8h。

（三）免疫印迹

在固相支持物上以抗原抗体的特异性反应检测混杂蛋白质中的特异成分。常用的方法有放射自显影、酶免疫测定及免疫荧光法等，其中以酶免疫测定法最常用，目前多以酶标抗体间接染色进行检测。

【材料】

洗膜液：含 0.5% Tween-20 的 PBS。

第一抗体溶液：用封闭液稀释第一抗体至 1:200 到 1:2000。

第二抗体溶液：用封闭液将酶标第二抗体稀释至 1:200 到 1:2000。

显色液：含 0.01% H_2O_2 和 0.05% 二氨基联苯胺（DAB）的 25mmol/L Tris、0.14mol/L NaCl 缓冲液（现配现用，放置不能超过 2h）。

【方法】

1. 封闭后的膜用含 0.5% Tween-20 的 PBS 洗 3 次，每次 15min。

2. 将 NC 膜置于小袋中，加入稀释的第一抗体，浸没 NC 膜，4℃ 过夜。

3. 取出膜置于平皿中，用含 0.5% Tween-20 的 PBS 中洗 3 次，每次 15min。

4. 再将 NC 膜置于小袋中，加入稀释的酶标二抗浸没 NC 膜，室温 1～2h。

5. 取出膜置于平皿中，用含 0.5% Tween-20 的 PBS 洗 3 次，每次 15min。

6. 显色　在抗原抗体反应过程中，配制显色液，过滤后备用。抗原抗体反应结束后，将膜置于显色液中显色，直至显色效果满意为止。

7. 显色后立即将膜置于双蒸水中终止反应。

【注意事项】

1. 蛋白质条带从凝胶原位电转移至 NC 膜时，一定要将 NC 膜一方接正极，否则蛋白质将丢失。

2. 电转移时间视蛋白质分子大小及电荷情况而不同，需预先摸索。

3. 电转移后一定要在封闭液中封闭，以消除背景。

4. 免疫印迹过程中抗体一定要浸没膜，并应不断翻动，保证反应均匀。

5. 显色液一定要现配现用，显色时间因抗体滴度而不同：滴度低时间长，反之时间短。显色满意后立即将膜转入双蒸水中漂洗终止反应。

（王晓燕　王　群）

实验十三 原位杂交

Exp13. In Situ Hybridization

【原理】

原位杂交技术由 Gall 和 Pardue 及 John 等于 1969 年建立，其基本原理是利用已知的标记核酸探针，通过碱基配对的原则，对组织或细胞内的未知核酸(DNA 或 RNA)进行定性定位的方法。一般应用的原位杂交有两种类型，与细胞内 DNA 的杂交和与细胞内 RNA 的杂交。这两种类型都是将标本固定在显微镜载玻片上，经杂交产生杂交信号后，在光学显微镜下就可观察到结果。

常用的原位杂交技术可分为放射性核素原位杂交技术和荧光原位杂交技术(fluorescence in situ hybridization, FISH)。它不仅可用于细胞学和组织学的研究，而且在研究病毒与肿瘤的关系上具重要作用。近几年来，由于染色体显带技术、荧光标记技术及电镜技术的发展和完善，原位杂交技术在生物学研究上展示了更加广阔的应用前景。

下面我们以甲醛固定、石蜡包埋的肝癌组织，地高辛标记的 HBV-DNA 探针为例，简单介绍原位杂交的方法。

【材料】

0.1%多聚赖氨酸。

4%多聚甲醛。

缓冲液 A:0.1mol/L Tris-HCl;0.15mol/L NaCl(pH7.4)。

封闭缓冲液:2%马血清;0.2%Tween-20;0.1mol/L Tris-HCl;0.15mol/L NaCl。

抗体缓冲液:1%马血清;0.3%Tween-20;0.1mol/L Tris-HCl;0.15mol/L NaCl。

底物缓冲液:100mmol/L Tirs-HCl;100mmol/L NaCl;50mmol/L $MgCl_2$,pH9.5。

NBT 和 X-磷酸盐显色剂。

HBV-DNA 探针。

碱性磷酸酶标记的地高辛(Dig)抗体。

【方法】

1. 载玻片的处理

(1) 新的载玻片放入硫酸中浸泡 24h，水冲洗 24h 后，用单蒸水浸泡 30min，酒精浸泡 30min，自然风干。

(2) 用 0.1%多聚赖氨酸涂布载玻片上，室内晾干备用。

2. 脱蜡

(1) 将 4～6μm 厚的肝癌石蜡切片置于处理好的载玻片上，60℃温箱过夜。

(2) 用二甲苯 37℃ 脱蜡 2 次，每次 2min，自然风干。

(3) 水化：将玻片放于 100%、90%、70%、50%、30% 的系列乙醇中各 3min，1×PBS 洗涤 5min。

(4) 0.2mol/L HCl 室温处理 10～20min，1×PBS 洗 5min，自然风干。

3. 杂交前预处理

(1) 每张切片滴加 50～100μl 20mg/ml 蛋白酶 K(proteinase K)，置湿盒内 37℃ 水浴 15min。

(2) 滴加 100μl 2mg/ml 甘氨酸-PBS，中止消化 5min，1×PBS 洗 5min。

(3) 4% 多聚甲醛(PEA)固定 5～10min，1×PBS 洗 5min。

(4) 脱水：30%、50%、70%、90%、100% 的系列乙醇脱水各 1min，自然风干。

4. 预杂交和杂交

(1) 每张切片滴加 100μl 预杂交液，置湿盒内 42℃ 水浴 1h。

(2) 倾去预杂交液，每张切片滴加 10μl 地高辛标记的 HBV DNA 探针(500ng/ml)，每张切片探针浓度为 5μg，置温盒内 96℃ 变性 5min，然后 42℃ 水浴过夜。

5. 漂洗

(1) 2×SSC 室温洗涤 10min。

(2) 0.1% SDS，0.1×SSC 于 37℃ 洗涤 30min。

(3) 0.1×SSC 于 37℃ 洗涤 10min。

(4) 缓冲液 A 洗 1min。

6. 显色

(1) 每张切片滴加 100μl 封闭缓冲液，37℃ 30min。

(2) 除去封闭缓冲液，每张切片滴加 30～50ml 酶标记抗体(5μg/ml)。

(3) 冲洗：用缓冲液 A 于室温洗 2 次，每次 5min。

(4) 于底物缓冲液中平衡 2min。

(5) 每张切片滴加 100μl 新配制的 NBT 和 X-磷酸盐显色液，暗室显色 6～24h(最长不超过 24h)。

(6) 于 TE 内中止反应 5min。

(7) 0.003% 吖啶橙水溶液，复染色 1～2min。

(8) 30%、50%、70%、90%、100% 的系列乙醇脱水，各 1min。

(9) 二甲苯透明两次，每次 1min，树脂封片，光镜观察拍照。

【注意事项】

1. 所用的载玻片一定要清洁，如有油污，则易导致杂交过程中组织标本脱片。

2. 显色时应闭光、静止。

(刘素侠　王　群)

实验十四　PCR-SSCP 技术

Exp 14. Single Strand Conformation Polymorphism

【原理】

单链构象多态性(single strand conformation polymorphism, SSCP)分析技术,是一种DNA单链凝胶电泳技术,它是根据形成不同构象的等长DNA单链,在中性非变性聚丙烯酰胺凝胶中电泳迁移率的变化,以检测分析基因变异的分析技术。如果DNA发生突变,则单链构象改变,电泳时表现出不同的带型,从而鉴定DNA的点突变。

聚丙烯凝胶电泳的凝胶介质分辨率很高,能分辨因点突变或者序列的变化使DNA分子的单链构象所发生的改变。用于SSCP分析的核酸片段小于200bp,检测灵敏度为70%。

本实验利用PCR-SSCP技术检测不同HBV感染者体内HBV基因的变异情况。

【材料】

电泳仪、垂直电泳槽、两块制胶玻璃板、玻璃板夹(夹层厚度1mm)、梳子、1%普通琼脂糖凝胶。

20%丙烯酰胺凝胶贮存液,亚甲双丙烯酰胺(Bis):丙烯酰胺(Acr)=1:19。

5×TBE浓贮存液1L含:54g Tris、27.5g 硼酸、20ml 0.5mol/L EDTA(pH8.0)。

10%过硫酸铵(AP)。

TEMED。

低离子强度(low ionic strength, LIS)缓冲液:含0.05%溴酚蓝,0.01%二甲苯青FF和10%蔗糖溶液。用LIS上样缓冲液处理的样品,由于溶液离子强度的降低可使核酸分子的解链温度和复性效率降低,因而能够提高单链生成率和分离效果。

固定液:10%乙醇+0.5%冰乙酸。

染色液:0.2% $AgNO_3$ 溶液。

显色液:1.5% NaOH+0.4%甲醛。

终止液:0.75% Na_2CO_3。

HBV病人血清标本。

【方法】

1. 目的基因片段的PCR扩增　利用HBV基因扩增试剂盒,对不同来源的血清标本进行PCR扩增,取小量进行普通琼脂糖凝胶电泳,观察产物的有无及DNA带亮度。

2. SSCP聚丙烯酰胺凝胶液的配制　用于SSCP分析的凝胶片段长度通常在500bp以

下，所以聚丙烯酰胺凝胶的浓度可选 8%，片段较小时选 12%。浓度为 8% 的凝胶液，其配制可按表实验 14-1 进行（所配液体体积可据玻璃板大小而定。以玻璃板大小 14cm × 11.5cm 为例，凝胶厚度 1mm，安装好的灌胶模具容量约为 11ml，所以所配凝胶液的量可按 12ml 计）。

表实验 14-1 8%凝胶中各成分的添加量

试 剂	20% BisAcr	H_2O	5×TBE	10% AP	TEMED
添加量	4.8 ml	4.8 ml	2.4 ml	84 μl	1.4 μl

3. 灌胶 将玻璃板装入橡胶架，固定于电泳槽中，插入合适的梳子，用 1% 琼脂糖凝胶密封两侧与底部。将凝胶溶液各成分按照表实验 14-1 中的顺序加入小烧杯，混匀后抽至 20ml 注射器中，带针头贴玻璃板内壁，从一角缓慢将液体注入两玻璃板空隙，以防气泡产生。室温聚合 1h，待梳齿下出现 Schlieren 线时，小心拔出梳子，立即用水冲洗孔格。在电泳槽上、下两槽中灌入 0.5×TBE 溶液，驱尽槽内气泡。

4. 样品处理 取 1～3μl PCR 产物（取量视产物浓度而定），加 20μl LIS-缓冲液，98℃ 放置 5min，取 5μl 上样。

5. 电泳分离 温度 15℃（室温较高时，可用冰袋敷于电泳槽外降温），电压 300V，电流 20mA，电泳时间 3h，片段较长时，可适当延长电泳时间。

6. 银染

(1) 取下凝胶板，DDW 水洗 2 次。

(2) 固定液固定 6min，水洗 2 次。

(3) 0.2% $AgNO_3$ 溶液染色 10min，水洗 3～5 次。

(4) 显色液显色 7min。注意甲醛易氧化，不要加得过早。

(5) 加入 0.75% Na_2CO_3 溶液终止显色。

【注意事项】

1. 丙烯酰胺是一种神经毒剂，应避免通过皮肤吸收。

2. 玻璃和垫条要彻底清洗干净，否则影响凝胶品质，加样时易形成气泡或漏出。

3. 冲洗凝胶孔格要充分，以免导致电泳带型不整齐。

4. 凝胶和电泳贮液槽中使用的 TBE 应是同时配制的，否则会干扰 DNA 的迁移。

5. 可根据电泳板的大小，配制相应体积的凝胶试剂。

6. PCR 产物在 LIS 条件下热变性，样品毋须立即转入冰浴中。一般在室温条件下，保存 4～5h 不会影响分析结果。

（于学慧 赵 丽）

实验十五　核酸电镜技术

Exp15. Electron Microscopic Technology Applied in Nuclear Acids

【原理】

许多碱性球蛋白如细胞色素 c 在低盐溶液或蒸馏水的表面容易发生变性形成不溶解的变性薄膜，并在液面上展开。如果所处的液面大小合适，这种薄膜即可扩展成单分子层而形成由伸展的肽链构成的蛋白质分子网。在展层溶液内，细胞色素 c 分子多肽链上的大量氨基酸碱性侧链基团（中性环境中细胞色素 c 带正电荷）与核酸分子链上的酸性基团（核酸带负电荷）以水合链结合（或靠静电吸引），核酸分子就相对稳固地固着在细胞色素 c 的周围，在展层时，随着细胞色素 c 在溶液表面展开，核酸也随之被拉开伸展了。用覆盖有碳膜的铜网膜面朝下在溶液表面沾一下，核酸即可吸附到碳膜上（核酸与碳膜的亲和力比核酸与下相液的亲和力大 10 倍），转移到碳膜上的核酸经染色，金属喷镀增加反差即可在电镜下观察。所用的碱性蛋白除细胞色素 c 外，还有糜蛋白酶、胰蛋白酶、核酸酶、溶菌酶等，但稳定性差，效果不如细胞色素 c。

【材料】

上相液（临用前配制）：10μl/ml 细胞色素 c；1μg/ml DNA；1mol/L NH_4Ac。

下相液：0～0.4mol/L NH_4NC。

三蒸水；滑石粉；滤纸；碳膜；铜网；90％乙醇；0.05mmol/L 醋酸。

电子显微镜；载玻片；TEFLON（聚乙氟乙烯）盘。

【方法】

1. 展层　将下相液倒入一个 TEFLON（聚乙氟乙烯）盘内（或者塑料培养皿内），使液面呈凸形。用 TEFLON 棒清洁液面后，将一干净载玻片一端浸入下液相内，一端搭在盘壁上，形成一个斜面。在下相液表面撒上一薄层滑石粉，用吸管吸取 50μl 上液相，沿距下液相表面 1 cm 的斜面上滴下，如展层良好，会看到形成的下液相将上液相表面滑石粉在瞬间内推向四周。待液面稳定后，用附有支持膜的铜网在液面上轻轻沾一下，用滤纸吸去多余的液体。

2. 核酸的电子染色　将沾有核酸的铜网放在一滴 0.05mmol/L 醋酸双氧铀液面上染色 30s，然后在 90％乙醇液滴上脱水 10s 之后用滤纸吸干，自然干燥。

3. 金属喷镀　在 1×10^5torrmmHg 的条件下，以 6.5°角，用钨丝（或钯-铱合金丝）加热

喷镀,样本台距喷射源的距离为 4.5cm,样本转动速度为 30r/min。

【注意事项】

展层时所用的器皿载玻片要绝对干净,不能有油脂。配制液体所用的水必须是三蒸水或去离子水。

（高立芬）

实验十六 表达谱 DNA 芯片样品的标记及杂交实验方案

Exp16. Sample Labeling and Hybridization Protocol of Expression Profile By DNA Chip

【原理】

目前用于表达谱研究的 DNA 的标记方法有直接标记方法、间接标记方法及体外转录的方法。间接标记法的基本原理是，以 RNA 样品为模板，在 RNA 指导 DNA 多聚酶（反转录酶）的催化下，以 oligo-dT 或随机引物指导，经反转录合成 cDNA，若在反转录的反应体系中加入氨基酰 dUTP（aadUTP），由于其结构类似于 dTTP，可以代替 dTTP 掺入到 cDNA 第一链中。掺入的 aadUTP 含有游离的氨基，可通过氨基与单功能的 Cy3 或 Cy5 染料进行偶联，从而对 cDNA 第一链进行荧光标记。与直接法相比，间接标记法大大提高了掺入的效率，减少了因 Cy3 和 Cy5 掺入效率的不同对结果的影响。荧光标记的样品、对照的 cDNA 探针与芯片上的基因片段进行竞争性杂交，探针的浓度越高，与芯片上的基因片段杂交的比例越高，经过激光共聚焦扫描，分别得到样品与对照的荧光强度的资料及比值的数据，经过进一步分析，即可得到样品与对照的某个基因的表达水平。根据芯片上所点基因的不同，可得到不同数量的差异表达基因，是一种高通量的检测基因表达的方法。可用于比较不同疾病状态与正常的基因表达差异，也可以比较不同处理之间的基因表达的差异。

【材料】

Microcon YM30 Centrifugal Filter Devices（Millipore ；Cat＃ 42410）。

QIAquick PCR Purification Kit（Qiagen; Cat＃ 28106）。

SuperscriptⅡ RT（200U/μl）（Life Technologies Cat＃18064-014）。

dNTP sets（Promega）。

5-（aminoallyl）-2′ deoxyuridine-5′ triphosphate（aadUTP）（Sigma; Cat＃ A0410），50 × aadUTP: 100mmol/L dATP，dCTP，dGTP，各 10μl，100mmol/L dTTP6μl，100mmol/L aadUTP 4μl，混合 −20℃ 保存

Cy3 NHS -easter（Amersham Pharmacia Cat＃Pa23001）荧光染料。

Cy5NHS -easter（Amersham Pharmacia;Cat＃PA 25001）荧光染料。

DMSO（Sigma），Human cot-1 DNA，polydT（20mers）（2.5mg/ml），羟胺（hydroxylamine）。

预杂交液：20×SSC 2.5μl，20%SDS 250μl，BSA 0.5g，加水至50ml，过滤除菌，-20℃保存。

杂交液：7.5μl 甲酰胺，3.75μl 20×SSC，1.5μlH_2O，0.75μl 2%SDS，1.5μl 1mg/ml human cot-1DNA。

总RNA或mRNA（要求OD_{260}/OD_{280}比值在1.8～2.0，提取方法见实验五）。

PCR仪或精密水浴锅，扫描仪，真空浓缩离心机，普通高速离心机，芯片杂交盒（Corning），染片缸，其他实验室常用试剂及耗材。

【方法】

1. cDNA第一链的合成

(1) 取一个1.5ml Eppendorf管，按下述加样：

oligo-dT(2.5μg/μl)	1μl
总RNA(10～20μg 或mRNA2μg)	×μl
0.05%DEPC-H_2O	补至15.5μl
总体积	15.5μl

混匀后，65℃水浴放置10min，置室温冷却10min。

(2) 另取一个1.5ml Eppendorf管，按下述加样：

5×st Strand buffer	6μl
0.1mol/L DTT	3μl
50×aadUTP mix	0.6μl
0.1%DEPC-H_2O	2.4μl
总体积	12μl

混匀后，42℃预热4min。

(3) 将RNA转移到反转录反应体系中，再加入

Superscript Ⅱ RT RNase H(200U/μl)　　2μl

RNA酶抑制剂　　0.5μl

0.05%DEPC-H_2O适量，使总体积为30μl。

42℃ 水浴内反转录2h，即可得到aadUTP标记的cDNA。

2. RNA水解　向上述反应体系中加1.6μl 5mol/LNaOH（终浓度0.25mol/L）及20μl 0.5mol/L EDTA，65℃水浴放置15min，使RNA水解，得到单链的cDNA。再加10μl 2mol/L HEPES(pH 8.0)中和反应。

3. DNA产物的纯化　以Microcon YM30纯化cDNA第一链，具体步骤如下：

向浓缩柱中加入450μl灭菌H_2O或TE，再加入反转录反应产物，12 000r/min，离心6～10min，弃去滤过产物，注意不要使柱的液体离干，重复两次，最后使柱中的液体保持在50μl左右。掉转柱子，将柱子置于一个新的1.5ml EP管中，13 000r/min离心，产物经真空浓缩离心机抽干。

4. 标记的cDNA与Cy5或Cy3染料的偶联　以4.5μl 0.1mol/L Na_2CO_3(pH 9.0)重悬cDNA，以4.5μl DMSO溶解荧光染料（NHS-ester-Cy3或Cy5）（1mg荧光染料以72μlDMSO

溶解,分装 16 个小管,-80℃保存,2 周内使用或真空抽干后 4℃保存)室温避光偶联 1h,加入 4.5μl 4mol/L 羟胺(0.28g,溶解到 1ml H_2O),室温 15min,终止反应。

5. 以 QIAQuick PCR Purification Kit 纯化 Cy5/Cy3 标记的偶联产物　将样品和对照的(Cy5 与 Cy3 标记)偶联产物合并,加 70μl H_2O。向 PCR 纯化柱中加 500μl buffer PB,13 000 r/min,离心30～60s,弃去滤过的液体,加 750μl buffer PE(应用液),13 000r/min,离心 30～60s,弃去滤过液体,重复一次。13 200r/min 空离 1min,干燥柱子。在柱中央加 30μl buffer EB,静止 1min,13 000r/min 离心 1min。重复洗脱一次。洗脱产物以真空浓缩离心机抽干。

6. 微阵列预处理

(1) 湿盒水合,80℃瞬间干燥 30s。

(2) 紫外交联,能量 190MJ。

(3) 0.1%SDS 浸泡 30s,水浸 30s。

(4) 沸水中变性 3min。

(5) 70%乙醇浸泡 1～2min。

(6) 500r/min 离心 5min。

7. 微阵列预杂交处理

(1) 加入预杂交缓冲液 42℃预杂交 45min。

(2) 将微阵列置于室温的蒸馏水中,使盖玻片脱落。

(3) 异丙醇洗涤(瞬间即可)后,置空气中干燥。

(4) 一小时之内使用微阵列。

8. 杂交

(1) 以 15μl 杂交液重悬荧光染料标记的 cDNA 探针,100℃水浴变性 2min,室温冷却 5min。

(2) 将盖玻片盖在微阵列上,从盖玻片的侧边加入探针。

(3) 42℃避光杂交 20～24h。

9. 杂交后洗涤(染片缸中进行,注意避光,插片时一定要轻)

(1) 2×SSC 0.1%SDS 洗涤 3min。

(2) 1×SSC 洗涤 2min。

(3) 0.2×SSC 洗涤 1min。

(4) 0.05×SSC 洗涤少于 10s。

(5) 500r/min 离心 3～4min 甩干,3h 内扫描。

10. 扫描　杂交后芯片的扫描,根据不同的标记方法采用不同的扫描仪进行扫描。荧光标记的探针采用激光共聚焦扫描仪扫描。扫描后的图像以 TIFF 格式保存,然后采用 Quantarray 进行分析。从 Quantarray 获得的数据是以文本格式存储的文件,包括 Cy5/Cy3 的比值,荧光强度及其他参数。这样得到的比值资料由于染料之间标记效率的不同还需要通过 Normalization 软件进行校对,最终获得的 Cy5/Cy3 的比值才能真正反应病理组织与正常组织中基因表达的差异,并通过聚类分析找出表达差异基因之间的相互关系,从而阐明基因的表达差异与疾病的发病机制之间的关系。

【注意事项】

1. RNA 的提取一般要求采用 Trizol 试剂提取，纯度 OD_{260}/OD_{280} 要求在 1.8～2.0 之间。

2. 荧光染料的保存很重要，分装后的染料应尽快的使用。

3. 荧光染料标记 cDNA 后，后续的操作尽量的避光。

4. 应对染料标记效率进行计算。计算公式如下：采用 $50\mu l$ 的比色杯，将标记后纯化的 cDNA 产物全部加入比色杯中，分别在 550nm 及 650nm 波长比色，计算各自的标记效率。

$$\text{pmol 核苷酸} = \frac{OD_{260} \times \text{体积} \times 37ng/\mu l \times 1000pg/ng}{324.5pg/pmol}$$

$$\text{pmolCy3} = \frac{OD_{550} \times \text{体积}(\mu l)}{0.15}$$

$$\text{pmolCy5} = \frac{OD_{650} \times \text{体积}(\mu l)}{0.25}$$

$$\text{核苷酸/染料比值} = \frac{\text{pmolcDNA}}{\text{pmol Cy 染料}}$$

要求每个样品染料结合率大于 200pmol，核苷酸/染料比值小于 50 时可用于杂交。

（韩金祥　高雪芹）

附　录

附录Ⅰ　常用玻璃器皿的处理

1. 玻璃器皿的处理　新购进来的玻璃器皿应先用肥皂水或洗洁精洗刷，后用自来水冲洗，再用1％～2％盐酸水溶液浸泡6～12h，以除去游离碱，用自来水冲洗干净后，再浸泡于硫酸清洁液中过夜，最后用自来水冲洗，玻璃瓶等器皿要注满水后倒出，如此反复12～15次，再用蒸馏水洗1～3次，自然晾干或者烘干。

培养细胞用的培养瓶、吸管、滴管等对清洁要求更高，用蒸馏水冲洗至少要5～8次，再用双蒸水冲洗5次以上。烘干，包装，高压，烘干，备用。

2. 硫酸清洁液的配制与注意事项

(1) 硫酸清洁液的配方如下：

成分	配方Ⅰ	配方Ⅱ
重铬酸钾(g)	80	100
自来水(ml)	1 000	750
粗浓硫酸(ml，工业用)	120	250

(2) 硫酸清洁液的配制与注意事项：

1) 配制时要注意安全，必要时戴上耐酸手套、穿上耐酸围裙，要特别注意保护面部和身体裸露的部位。

2) 先将重铬酸钾溶解于自来水中(可稍微加热)，待冷至室温后徐徐加入浓硫酸，不然容易爆沸造成意外事故。

3) 清洁液有很强的酸性和氧化能力，为防止器皿内残留的大量有机物与清洁液中的铬酸盐结合而迅速破坏失效，因此要求各类器皿需先用肥皂水或清洁液冲洗晾干后，再放入清洁液中。

4) 若玻皿内附有Hg^{2+}、Pb^{2+}、Ba^{2+}两价离子时，要先用稀盐酸或稀硝酸处理后再浸入硫酸清洁液。

5) 新配制的清洁液为红褐色，经反复使用会变成绿色，表明已经失效，应弃去不用。

3. 玻璃制品和塑料制品的硅烷化处理

(1) 把待硅烷化的物品放入一较大、清洁的玻璃干燥器内。

(2) 干燥器内放置一个小烧杯，内装1ml二氯二甲硅烷。

注意：二氯二甲硅烷有毒性和挥发性、高度易燃，务必在化学通风橱中操作。

(3) 通过接液瓶将干燥器与真空泵相连，开启真空系统抽气至二氯二甲硅烷开始沸腾，立即切断真空泵与干燥器之间的连接管道，关闭真空泵。干燥器内应维持真空状态，一旦二氯二甲硅烷开始沸腾，务必马上关闭真空泵，否则挥发剂将被抽入真空泵而破坏真空密封。

(4) 待二氯二甲硅烷挥发至尽(1～2h)，在化学通风橱中打开干燥器，让二氯二甲硅烷烟雾分散后，取出玻璃或塑料制品。玻璃制品和玻璃棉使用前应于180℃烘烤2h。塑料制品不应高压灭菌，但使用前须用水彻底冲洗

(5) 大件玻璃制品的硅烷化是把物品用氯仿或庚烷配制的5%二氯二甲硅烷浸泡或漂洗,有机溶剂挥发时,二氯二甲硅烷即沉积在玻璃制品上,用前应用水反复冲洗多次或于180℃烘烤2小时。

附录Ⅱ　常用培养基和溶液

1. 常用培养基、溶液

(1) LB(Luria - Bertani)培养基:Bacto-tryptone10g,Bacto-yeast extract 5g,NaCl 10g,溶于800ml双蒸水内,用NaOH调pH至7.5定容至1L,分装高压灭菌。

(2) B固体培养基:按琼脂:LB液体培养基(W/V)=1.5%～2.0%配成。

(3) 1mol/L Tris:溶解121.19g Tris碱,于800ml DDW内,加浓盐酸调所需pH。

pH7.4约需浓HCl 70ml,pH7.6约需浓HCl 60ml,pH 7.4约需浓HCl 42ml,在最终调pH之前,使溶液冷至室温,定容体积为1L。分装,高压灭菌。若溶液呈黄色时,弃之换用试剂纯Tris再配制新溶液。

(4) 0.5mol/L EDTA(pH8.0):称186.1g$EDTA\text{-}Na_2 \cdot H_2O$溶于800mlDDW中,振荡混匀使溶,用NaOH(20gNaOH颗粒)调pH至8.0。调体积至1L,分装,高压灭菌。

(5) 5mol/L NaCl:溶解292.2g NaCl于800mlDDW中,调体积至1L,分装,高压灭菌。

(6) 1mol/L $MgCl_2$:溶解203.3g $MgCl_2 \cdot 6H_2O$于800mlDDW中,调体积1L,分装,高压灭菌。

(7) 3mol/L NaAC(pH5.2):溶解408.1g $NaAC \cdot 3H_2O$ ($CH_3CO_2Na \cdot 3H_2O$)于800ml DDW中,用冰醋酸调pH至5.2,调体积至1L,分装,高压灭菌。

(8) 1mol/L DTT:溶解3.09g二巯苏糖醇(DTT)于20ml 0.mol/L NaAC(pH5.2),过滤除菌,分成1ml/管,-20℃保存。DTT或含DTT的溶液不能高压灭菌。

(9) EB 10mg/ml:加1g溴化乙锭(EB)到100ml DDW中,磁力搅拌几小时以确保染料溶解,置暗瓶内,瓶外包上铝箔纸,4℃保存。EB是一种强锈变剂,操作者要戴手套注意防护自身,更应注意避免污染环境。

(10) 10% SDS:在900ml水中溶解100g电泳级SDS,加热至68℃助溶,加入几滴浓盐酸调pH为7.2,加水定容至1L,分装备用。SDS溶液无需灭菌。

(11) 20×SSC:在800ml水中溶解173.5g NaCl和88.2g枸橼酸钠,加入数滴10mol NaOH溶液调pH至7.0,加水定容至1L,分装后高压灭菌。

(12) X-gal:用二甲基甲酰胺溶解X-gal,配制成120mg/ml(W/V)的贮存液,保存于一玻璃管或聚丙烯管中,用铝箔纸封裹好以防光照,贮存-20℃。X-gal溶液无须过滤除菌。

2. 抗生素溶液的配制

(1) 氨苄西林(AP)保存液:25mg/ml(W/V)氨苄西林钠盐溶液,过滤除菌。分装,-20℃保存。工作浓度为35～50μg/ml。

(2) 氯霉素(CM)保存液:为34mg/ml(W/V)的100%乙醇溶液,于-20℃保存。作用浓度:扩增质粒为170μg/ml;筛选抗性菌用30μg/ml。

(3) TC保存液:为12.5mg/ml(W/V)盐酸四环素的乙醇/水(50%,V/V)溶液。作用

浓度为 12.5μg/ml。

注:四环素对光敏感,应把含有四环素的溶液与平板保存在暗处;镁离子是四环素的拮抗剂,应使用无镁盐的培养基(如 LB)筛选四环素抗性菌。

3. 常用电泳缓冲液

(1) TAE 缓冲液:

1) 工作液(1×):0.04mol/L Tris-HAC, 0.001mol/L EDTA。

2) 浓缩保存液(50×):每升含 Tris 碱 242g,冰醋酸 57.1ml,0.5mol/L EDTA(pH8.0) 100ml。

(2) TPE 缓冲液:

1) 工作液(1×):0.09mol/L Tris-H_2PO_4,0.002mol/L EDTA。

2) 浓缩保存液(10×):每升含 Tris 碱 108g,85%磷酸(1.679mg/ml)15.1ml, 0.5mol/L EDTA(pH8.0)40ml。

(3) TBE 缓冲液:

1) 工作液(1×):0.045mol/L tris-硼酸,0.001mol/L EDTA。

2) 浓缩保存液(5×):每升含 Tris 碱 54g,硼酸 27.5g,0.05mol/L EDTA(pH8.0) 20ml。

(4) Tris-甘氨酸缓冲液:

1) 工作液(1×): 25mmol/L Tris, 250mmol/L 甘氨酸,0.1%SDS。

2) 浓缩保存液(5×):15.1g Tris 碱,94g 甘氨酸(电泳级,pH8.3),50ml 10%SDS(电泳级)。

4. 凝胶加样缓冲液

(1) Ⅰ型 6×缓冲液:0.25%溴酚蓝(PBS),0.25%二甲苯氰醇 (xylene cyanol),40%(W/V)蔗糖,水溶液于 4℃保存。

(2) Ⅱ型 6×缓冲液:0.25%溴酚蓝(PBS),0.25%二甲苯氰醇 (xylene cyanol),0.25% Ficoll(type 400),水溶液于室温保存。

(3) 2×SDS 凝胶加样缓冲液:100mmol/L Tris-HCl(6.8),200mmol/L DTT(二硫苏糖醇),4% 溴酚蓝,20% 甘油。

注:此缓冲液为不含二硫苏糖醇的 2×SDS。该凝胶加样缓冲液可保存于室温,临用前取 1mol 的二硫苏糖醇储存液,加于上述缓冲液中。

5. 限制性内切酶解缓冲液

缓冲液	NaCl (mmol/L)	Tris-HCl (mmol/LpH7.5)	$MgCl_2$ (mmol/L)	DTT (mmol/L)	BSA (μg/ml)
低盐	0	10	10	1	50
中盐	50	10	10	1	50
高盐	100	50	10	1	50

6. 凝胶浓度与 DNA 的分离范围

(1) 琼脂糖凝胶浓度与线形 DNA 的分离范围

琼脂糖凝胶的浓度/(%)	线形 DNA 长度/(bp)
0.5	1000～30 000
0.7	800～12 000
1.0	500～10 000
1.2	400～7000
1.5	200～3000
2.0	50～2000

(2) 聚丙烯酰胺凝胶电泳分离的范围

丙烯酰铵浓度/(%)	分离核苷酸的范围/(bp)
3.5	100～1 000
5.0	80～500
8.0	60～400
12.0	40～200
20.0	10～100

7. 不同浓度聚丙烯酰胺凝胶的配制

试　剂	聚丙烯酰铵浓度/(%)				
	3.5	5.0	8.0	12	20
丙烯酰铵溶液(ml)	11.6	16.6	26.6	40.0	66.6
10%过硫酸胺(ml)	0.5	0.5	0.5	0.5	0.5
10×TBE(ml)	10	10	10	10	10
总体积(ml)	100	100	100	100	100

附录Ⅲ　基因工程常用数据和换算计算公式

1. 基因组

R17 噬菌体：3.6kb，含 4 个基因(2000 年)。

E.coli(PB51)噬菌体:250kb,含 240 多个基因(2000 年)。

E.coli(K-12 MG1655)：4.64Mb，含 4288 个基因，其中 1/3 基因的功能尚不清楚(2001 年)。

结核杆菌:4.41Mb(2001 年)。

酵母:6000 个基因(2001 年)。

果蝇:13000 个基因(2002 年)。

线虫:18999 个基因(2002 年)。

拟南芥菜:26000 个基因(2002 年)。

人:31000 个基因,32 亿对碱基(2002 年)。

2. 常用核酸蛋白换算数据

(1) 重量换算:

$1\mu g = 10^{-6} g$

$1ng = 10^{-9} g$

$1pg = 10^{-12} g$

$1fg = 10^{-15} g$

(2) 分光光度换算:

$1A_{260}$DS DNA = 50μg/ml

$1A_{260}$SS DNA = 33μg/ml

$1A_{260}$DS RNA = 40μg/ml

(3) DNA:

平均碱基对的相对分子质量大约是 650 道尔顿

每摩尔单碱基对平均重约 660g

10^6 道尔顿的双链 DNA 大约是 0.5μm 长

10^6 道尔顿的双链 DNA 大约是 1.5kb

(4) DNA 摩尔换算:

1μg 1000bpDNA = 1.52pmol = 3.03pmol 末端浓度

1μg PBR322DNA = 0.36pmolDNA

1pmol 1000bpDNA = 0.66μg

1pmol PBR322DNA = 2.8μg

1kb DS DNA(钠盐) = 6.6×10^5 道尔顿

1kb SS DNA(钠盐) = 3.3×10^5 道尔顿

1kb SS RNA(钠盐) = 3.4×10^5 道尔顿

(5) 脱氧单核苷酸的平均相对分子质量 = 330g/mol:

1kb DNA = 330 个氨基酸编码容量

333 个氨基酸 = 3.7×10^4 道尔顿

(6) 蛋白摩尔换算:

100pmol 分子量 100 000 蛋白质 = 10μg

100pmol 分子量 50 000 蛋白质 = 5μg

100pmol 分子量 10 000 蛋白质 = 1μg

氨基酸的平均分子量 = 126.7 道尔顿

(7) 蛋白质/DNA 换算:

1000 分子量蛋白质 = 270bp DNA

30 000 分子量蛋白质 = 810bp DNA

50 000 分子量蛋白质 = 1.35kp DNA

100 000 分子量蛋白质 = 2.70kp DNA

注:1kb DNA = 333 个氨基酸编码容量 = 3.7×10^4 道尔顿蛋白质。

dA = dA 的数目, dC = dC 的数目, dG = dG 的数目, dT = dT 的数目。

3. 常用的计算公式:

(1) DNA 的摩尔浓度:

1) DNA 片段的分子量克/摩尔 DNA = 碱基对数 × 单碱基对平均重量

2) DNA 片段 5′-末端或 3′-末端的摩尔浓度 = DNA 克数/DNA 分子量 = 摩尔浓度 × 2

3) 环状 DNA 分子被限制性内切酶切割后的末端摩尔浓度 = DNA 摩尔浓度 × 切割位点数 × 2

4) 线性 DNA 分子被限制性内切酶切割后的末端摩尔浓度 = DNA 摩尔浓度 × 切割位点数 × 2 + 2

(2) 摩尔浓度和当量浓度的计算

1) 摩尔浓度:是 1L 溶液中含溶质的摩尔数,或 1ml 溶液中含溶质的毫摩尔数,以mol/L 表示,其表达式为: $mol/L = \frac{n}{V_L} = \frac{mn}{V_{ml}}$

式中:n,溶质的摩尔数;mn,溶质的毫摩尔数;V_L,溶质的体积升数;V_{ml},溶液的体积毫升数。

注: $n = \frac{W(g)}{fw(g)/M}$; W:物质的质量;fw:1 摩尔质量;M:分子量。

2) 当量浓度:是 1L 溶液中所含溶质的克当量数,或 1ml 溶液中所含溶质的毫克当量数,用 N 表示,其表达式为: $N = \frac{n}{V_L} = \frac{mn}{V_{ml}}$

式中:n,溶质的克当量数;mn,溶质的毫克当量数;V_L,溶液的体积升数;V_{ml},溶液的体积毫升数。

注: $n = \frac{W(g)}{E}$　W:物质质量　E:克当量

4. 常用的 DNA 碱基长度标准参照物

λDNA/*Hin*dⅢ	λDNA/*Hin*dⅢ + *Eco*RⅠ		pBR322/BstnⅠ	Φχ174/HaeⅢ		Gene Ruler 100bp DNA Ladder Plus	
23130	21227	1375	1857	1353	194	3000	600
9416	5148	974	1060	1078	118	2000	500
6557	4973	831	929	872		1500	400
4361	4268	564	383	603		1200	300
2322	3530	125	121	310		1031	200
2027	2027		13	281		900	100
564	1904			271		800	
125	1584			234		700	
(0.12~23.1 kb)	(0.12~21.2 kb)		(0.01~1.85 kb)	(0.10~1.35 kb)		(0.10~3.00 kb)	

5. 常用蛋白质相对分子质量标准参照物

	蛋白质	相对分子质量
	肌球蛋白	212 000
高	β-半乳糖苷酶	116 000
分	磷酸化酶 B	97 000
子	牛血白蛋白	66 200
量	过氧化氢酶	57 000
	醛缩酶	40 000
	磷酸化酶 B	97 400
	牛血白蛋白	66 200
中	谷氨酸脱氢酶	55 000
分	卵白蛋白	42 700
子	醛缩酶	40 000
量	碳酸苷酶	31 000
	大豆抑肽酶	21 000
	溶菌酶	14 400
	碳酸苷酶	31 000
	大豆抑肽酶	21 000
低	马心肌球蛋白	16 900
分	溶菌酶	14 400
子	肌球蛋白(F1)	8 100
量	肌球蛋白(F2)	6 200
	肌球蛋白(F3)	2 500

6. 同位素资料:

放射性测量单位:伦琴(r)是使 1g 空气产生 1.61×10^{12}离子对的 X 或 γ 射线的照射剂量。

R/0.87 = 1rad

$1Ci = 3.7\times10^{10}$原子量

$1\mu Ci = 3.7\times10^{4}$ 个核衰变/s = 1.2×10^{6} 个核衰变/min

阿伏伽德罗数 = 6×10^{23}

7. 典型大肠杆菌的重量:10^{-12}g/湿重/细胞;10^{-13}g 蛋白质/细胞;2×10^{-13}干重/细胞。

$1.0\times OD_{55}$(Zeiss)大约 = 1×10^{-9}mol。

8. 几个受体菌的表型:

(1) C600(*E. Coli*):来源于 K12(*E. Coli*),经人工改造后成为(F^-),(λ^-),thr-1,leuB6,LacY1,tpnA21,supE44. 是消除了 F 因子和 λ 噬菌体的多重缺陷型菌株。

(2) HB101(*E. Coli*):来源于 E.Coli,是 F ,hsdS20,(rB 、MB),supE44,recA13,are-14,proA2,lacY1,galK2,rpsL(Str)xyl-5,mlt-1,supE44,λ^- 的菌株。

(3) JM109(*E. Coli*):recA1,supE44 ,endA1,hsdR17(rK ,mK^+),gyrA96,mtl-1, relA1, thiΔ(lac^- proAB),λ^-,[F′traD36, $proAB^+$,$lag1^qZ\Delta M15$],γ(DE3)菌株。

(曹英林　张　蘋　郭　春)

附录Ⅳ 人类基因组测序参加单位索引(HGSI Participants)

BCM-HGSC	Baylor College of Medicine Human Genome Sequencing Center	US
CMG-WCH	Dept. of Cytogenetics and Molecular Genetics at the Women's And Children's . Hospital	Australian
GBF	GBF-Dept. of Genome Analysis, Braunschweig	Germany
GENOSCOPE	GENOSCOPE National Sequencing Center	France
GTC	Genome Therapeutics Corp.	US
HGGIGCAS	Human Genome Center , Institute of Genetics , Chinese Academy of Sciences	China
IMBB/ FORTH	Institute of Molec. Biology & Biotechnology/Foundation for Research and Technology Hellas	Greece
Jena	The Institute for Molecular Biology Jena	Germany
JGI	Joint Genome Initiative (LBNL/LANL/LLNL)	US
Keio	Keio University School of Medicine sequencing Team	Japan
MD-ACC	MD-Anderson Cancer Center	US
MMGC	Medical and Molecular Genetics Center	Spain
MOIMG-LH	Max-Planck-Institute for Molecular Genetics	Germany
RIKEN-GSC	RIKEN Genomic Sciences Center (formerly Tokyo , JST-Kitasato/U. Tokyo Sequencing Team)	Japan
Sanger	The Sanger Center	UK
SHGC	Stanford Human Genome Center	US
TIGR	The Institute for Genomic Research	US
UO-ACGT	University of Oklahoma	US
UTSW	University of Texas Southwest	US
UWGC	University of Washington Genome Center	US
UW-MSG	University of Washington Multimegabase Sequencing Group	US
WshU	Washington University School of Medicine Genome Sequencing Center	US
WI/MIT	Whitehead Institute /MIT Genome Sequencing Center	US
YMGC	National Yang Ming University Genome Research Center	Taiwan

(孙汶生 刘玉刚)

附录Ⅴ 专业词汇(英中对照)

anneal 退火
Aspergillus oryzae 稻谷曲霉
adaptor 普适子
anti-sense RNA 反义 RNA
aminopterin 氨基蝶呤
asymmetric PCR 不对称 PCR
allele specific oligo nucleotide probe ASO 探针(等位特异寡核苷酸探针)
arraying and replicating device ,ARD 阵列复制器
arrayer 阵列点样机
amber 琥珀
association analysis 关联分析
avian myeloblastosis virus, AMV 鸟类骨髓母细胞瘤病毒
bacterial alkaline phosphatase,BAP 细菌碱性磷酸酶
bacteriophageλ λ噬菌体
baculovirus 杆状病毒
biochip 生物芯片
bispecific antibody, BsAb 双特异性抗体
bifunctional antibody, BfAb 双功能抗体
clinical phenotype 临床表型
coding SNP,cSNP 编码区 SNP
chloramphenicol acetyltransferase, CAT 氯霉素酰基转移酶
cis-acting factor 顺式调控因子
compatibility 相容性
C value paradox C值矛盾
cellular genome 细胞基因组
chromatin 染色质
chromosome 染色体
coding strand (sense strand) 编码链(有义链)
consensus sequence 一致序列
covalent close circle, CCC 共价闭环
clone 克隆
cohesive ends (stick ends) 黏性末端
competent 感受态
complementary DNA,cDNA 互补 DNA
complete digestion 完全酶切消化作用
calf intestinal alkaline phosphatase,CAP 小牛肠碱性磷酸酶
cloning vectors 克隆载体
cosmid 黏性质粒
cohensive-end site COS 位点, 黏性末端位点
capping 带帽
cDNA library cDNA 文库
competent cell 感受态
core promoter element 核心启动子元件
cytomegalo virus, CMV 巨细胞病毒
co-transfection 共转染
code bias 编码偏爱性
contiguous stacking hybridization ,CSH 邻堆杂交技术
cyclin-dependent kinas es,CDKs Cyclin 依赖型激酶
clustering analysis 聚类分析
cyclone storage phosphor system 磷感屏成像系统
charge-coupled devices ,CCD 电荷偶合装置
chromosome translocation 染色体转位方式
cis-acting element 顺式调控元件
core sequence,module,motif 核心序列
CDR grafting antibody CDR 移植抗体
combinatorial antibody library 组合抗体库
differential display RT-PCR , DDRT-PCR 差异显示反转录 PCR

denature　变性
DNA chimera　DNA 嵌合体
DNA engineering　DNA 工程
distant cleavage　远距离裂解酶
DNA recombination　DNA 重组
DNA vaccine　DNA 疫苗
dehancer　衰减子
dihydrofolate reductase, dhfr　二氢叶酸还原酶
dot or slot blot　斑点及狭缝印迹法
DNA sequencing　DNA 测序
direct sequencing,DS　直接测序
DNA chips　DNA 芯片
DNase Ⅰ　核酸内切酶Ⅰ
DNAase Ⅰ hypersensitive site　DNase Ⅰ 超敏感
DNase Ⅰ hyperseneitivity assay　DNase Ⅰ 超敏感分析
differential library　差异显示文库
dependovirus　依赖性病毒
diabodies　双体
DNA polymorphism　DNA 多态性
enhancer　增强子
expression vector　表达载体
euchromatin　常染色质
expression-library immunization ,ELI　表达文库免疫技术
E. coli DNA pol Ⅰ Klenow fragment　大肠杆菌 DNA 聚合酶Ⅰ的 Klenow 片段
eukaryotic virus vector　真核细胞病毒载体
electroporation　电穿孔
expression vector　表达载体
exon　外显子
etoposide　鬼臼亚乙苷
epithelial glycoprotein　上皮糖蛋白
embryonic stem cells,ES　细胞胚胎干细胞
functional genomics　功能基因组学
filamentous phage　单链丝状噬菌体
framework region, FR　骨架区
flanking sequence　侧翼序列
flush(blunt) end　平齐末端
foreign DNA　外源基因
genetic engineering antibody　基因工程抗体
gene therapy　基因治疗
gel retardation assasy　凝胶滞留法
gene amplification　基因放大
gene chips　基因芯片
gene bank　基因库
gene　基因
gene immune　基因免疫
gene expression　基因表达
gene cloning　基因克隆
gene library　基因文库
gene manipulation　基因操作
gene recombination　基因重组
genetic engineering　基因工程
genomic library　基因组文库
genomic DNA library　基因组 DNA 文库
gene expression in prokaryotic cells　原核表达系统
gene expression in eukaryotic cells　真核基因表达系统
gutless vector　无病毒载体
genomic amplification with transcript sequeucing ,GAWTS　基因组扩增转录同步测序法
gene clustering　基因聚类分析
germ cells　性细胞
germline　胚系
gene knockout　基因敲除
gene targeting　基因打靶
gene knockin　基因敲入
human genome project, HGP　人类基因组计划
human anti-mouse antibody, HAMA　人抗

鼠抗体
humanized antibody 人源化抗体
house-keeping gene 看家基因
hybrid clustering approach 混合聚类法
heat shock protein, HSP 热休克蛋白
homopolymeric tail 同聚物尾巴
high repetitive sequences 高度重复序列
hybridization 杂交
heterochromatin 异染色质
heterogenuous nuclear RNA, hnRNA 核内异质 RNA
human genome 人类基因组
hybrid site 杂种位点
hypoxanthine 次黄嘌呤
hairpin structures 发夹结构
inner cell masses 内细胞团
isopropyl-β-D-thiogalactoside, IPTG 异丙基-β-D-硫代半乳糖苷
incompatibility 不相容性
immunostimulatory DNA sequence, ISS 免疫刺激 DNA 序列
isoschizomers 异源同工酶
isocaudamer 同尾酶
immunostimulatory DNA sequence ,ISS 免疫刺激 DNA 序列
interrupted or discontinuous gene 不连续基因
intervening sequence, IVS 插入序列
insertional inactivation 插入失活
intergenic sequence,IS 间隔顺序
intron 内含子
infection 感染
inverse PCR,IP-CR 反向 PCR
linker 接头
long terminal repeat sequence,LTR 长末端重复序列
liposome 脂质体
ligase 连接酶
lytical pathway 溶菌途径
lysogenic pathway 溶源途径
LacZ-α complementation peptide, ZAP LacZ-α 互补肽
linker 连接子
long PCR 长距离 PCR
linkage analysis 连锁分析
microinjection 显微注射法
monoclonal antibody, McAb 单克隆抗体
mobility shift assasy 迁移改变法
multiple cloning site,MCS 多克隆酶切位点
Moloney murine leukemia virus, M-MLV 莫洛尼鼠白血病病毒
mitochondrial DNA, mtDNA 线粒体 DNA
molecular cloning 分子克隆
multiple cloning sites, MCS 多克隆位点
modification nuclease 修饰性核酸酶
moderate repetitive sequences 中度重复序列
methotrexate, MTX 甲氨蝶呤
mini-Ad 微型腺病毒载体
multiplex PCR 多重 PCR
miniantibodies 微抗体
mini-satellites 微卫星 DNA
northern blot Northern 印迹
nucleic acid vaccine, NAV 核酸疫苗
nucleic acid 核酸
nucleoid 类核
nucleosome 核小体
nucleotide acid 核苷酸
nick translation 缺口平移
nucleic Acid Hybridization 核酸杂交
neural network approach 神经网络方法
normalization 标准化处理
nuclear run on transcription 进行中的核转录分析
native antibody library 天然抗体库

non-synonymous cSNP 非同义 cSNP
open reading frame, ORF 开放阅读框架
open circle, OC 开环
open reading frame, ORF 开放读码框架
osteoporosis 骨质疏松症
osteoclasts 破骨细胞
polyclonal antibody, PcAb 多克隆抗体
polynucleotide 多核苷酸
posttranslational processing 翻译后加工
polymerase chain reaction, PCR 聚合酶链反应
partial digestion 局部酶切消化
phage M13 M13 噬菌体
particle bombardment 微粒子轰击法
promoter 启动子
physical map 物理图谱
promoter for repressor establishment 阻抑物启动子 PRE
polylinker 多聚连接子
polymerase chain reaction, PCR 聚合酶链反应
primer dimer 引物二聚体或引物二倍体
primer walking 引物延伸法
pairwise average-linkage cluster analysis 配对平均连锁聚类分析
piezoelectric printing 压电打印法
phage antibody library 噬菌体抗体库
phage λ λ 噬菌体
phagemid 噬菌粒
plaque lift 菌斑印迹
phage display 噬菌体表面呈现技术
photomultiplier tube, PMT 光电倍增管
pharmacogenomics 药物基因组学
positive and negative selection, PNS 正负筛选法
quantitative PCR PCR 定量
reshaped antibody, RAb 改型抗体
restrictive enzyme fragment long polymorphogenesis ,RFLP 限制性酶片段长度多态性分析法
renature 复性
repetitive sequence 重复序列
RNA splicing RNA 剪接
recombinant DNA 重组 DNA
restriction endonuclease, RE 限制性核酸内切酶
relaxed plasmid 松弛型质粒
Ratio analysis Ratio 分析
reportor gene 报告基因
restriction endonuclease igestion and subcloning 限制性酶切-亚克隆法
reverse transcriptase 反转录酶
revert transcript polymerase chain reaction , RT-PCR 反转录 PCR
restriction enzyme, RE 限制性内切酶
recombinant antigen vaccine 重组抗原疫苗
recombinant 重组体
restriction map 限制性酶图谱
rous sarcoma virus, RSV 肉瘤病毒
real-time quantitative PCR 实时荧光定量 PCR
RNase RNA 酶
recombinant PCR, RP-CR 重组 PCR
RNA ploymerase, RNA pol RNA 聚合酶
Rnase mapping RNase 作图法
repertoire antibody 抗体库
restriction fragment length polymorphisms, RFLPs 限制性片段长度多态性
RNA interference, RNAi RNA 干预技术
short tandem repeats, STRs 短串联重复序列
single chain of Fv, ScFv 单链抗体
somatic cells 体细胞
silencer 沉寂子
splicing 剪切
sceening 筛选

single nucleotide polymorphisms, SNP 单核苷酸多态性
star activity 星号活性
stringent plasmid 严紧型质粒
single copy sequences 单拷贝序列
shuttle vector 穿梭质粒
single strand conformation polymorphism analysis of polymerase chain reaction products, PCR-SSCP 聚合酶链反应－单链构象多态性
Southern blot Southern 印迹
shotgun sequencing 鸟枪法
sequencing by hybridization, SBH 杂交测序
scanning tunneling microscopy, STM 扫描隧道显微术
sequencing by hybridization, SBH 杂交测序
Saccharomvces cerevisiae 酿酒酵母
self organizing maps 自组织图
splicing 拼接
S_1 -mapping 核酸酶 S_1 作图法
subtracted library 扣除文库
small-satellites 小卫星 DNA
synonymous cSNP 同义 cSNP
serological analysis of recombinant cDNA expression libraries, SEREX 重组 cDNA 表达文库的血清学分析法
simian vacuolating virus40, SV40 绿猴空泡病毒
semi-conservative replication 半保留复制
split gene 断裂基因
structural gene 结构基因
supercoil, SC 超螺旋
telomerase 端粒酶
telomere 端粒
temperature melting, T_m 解链温度
transformation 转化
transgenic animal 转基因动物
template strand (antisense strand) 模板链(反义链)
Taq DNA polymerase *Taq* DNA 聚合酶
terminal deoxynucleotidyl transferase 末端脱氧核苷酸转移酶
terminal transferase 末端转移酶
transcription vector 转录载体
transcription in early stage 早期转录
transcription in delayed early stage 迟早期转录
transcription in late stage 晚期转录
transformation 转化
transfection 转染
terminator 终止子
thrombin 凝血酶
trans-acting factor 反式作用元件
thymidine 胸苷
terminal repeat 末端重复区
transfer vector 转移载体
transfer vector plasmid 转移载体质粒
tetramethylrhodamine 四甲基若丹明
trans-acting factor 反式作用因子
transition 转换
transversion 颠换
upstream promoter elements 上游启动子元件
unique sequence 单一序列
unsupervised Clustering 非监督聚类法
variable number tandem repeats, VNTRs 可变数目串联重复序列
vector 载体
western blot 免疫印迹
yeast plasmid 酵母质粒
yeast inregrating plasmid, YIp 酵母整合型质粒
yeast replicating plasmid, YRp 酵母复制型质粒

yeast episomal plasmid, YEp 酵母附加子型载体

yeast artifical chromosome, YAC 酵母人工染色体

(郭 春)

附录Ⅵ 常用专业名词与解释(英文)

Adaptor A synthetic, double-stranded oligonucleotide used to attach sticky ends to a blunt-ended molecule.

Adenovirus An animal virus, derivatives of which have been used to clone genes in mammalian cells.

Affinity chromatography A chromatography method that makes use of a ligand that binds a specific protein and which can therefore be used to aid purification of that protein.

Agrobacterium tumefaciens The soil bacterium which, when containing the Ti plasmid, is able to form crown galls on a number of dicotyledonous plant species.

Ancient DNA Preserved DNA from an archaeological or fossil specimen.

Annealing Attachment of an oligonucleotide to a single-stranded DNA molecule by hybridization.

Antisense RNA An RNA molecule that is the reverse complement of a naturally occurring mRNA, and which can be used to prevent translation of that mRNA in a transformed cell.

Antisense technology The use in genetic engineering of a gene coding for an antisense RNA.

Autoradiography A method of detecing radioactively labelled molecules through exposure of an X-ray-sensitive photographic film.

Auxotroph A mutant microorganism that grows only when supplied with a nutriend not required by the wild-type.

Avidin A protein that has a high affinity for biotin and is used in a detection system for biotinylated probes.

Bacterial artificial chromosome (BAC) A cloning vector based on the F plasmid, used for cloning relatively large fragments of DNA in *E. coli*.

Bacteriophage or **phage** A virus whose host is a bacterium. Bacteriophage DNA molecules are often used as cloning vectors.

Baculovirus A virus that has been used as a cloning vector for the production of recombinant protein in insect cells.

Batch culture Growth of bacteria in a fixed volume of liquid medium in a closed vessel, with no additions or removals made during the period of incubation.

Bioinformatics The use of computer methods in studies of genomes.

Biolistics A means of introducing DNA into cells that involves bombardment with high velocity microprojectiles coated with DNA.

Biological containment One of the precautionary measures taken to prevent the replication of

recombinant DNA molecules in microorganisms in the natural environment. Biological containment involves the use of vectors and host organisms that have been modified so that they will not survive outside of the laboratory.

Biotechnology The use of biological processes in industry and technology.

Biotin A molecule that can be incorporate into dUTP and used as a non-radioactive label for a DNA probe.

BLAST An algorithm frequently used in homology searching.

Blunt end or **flush end** An end of a DNA molecule at which both strands terminate at the same nucleotide position with no single-stranded extension.

Broad host range plasmid A plasmid that can replicate in a variety of host species.

Broth culture Growth of microorganisms in a liquid medium.

Buoyant density The density possessed by a molecule or particle when suspended in an aqueous salt or sugar solution.

Candidate gene A gene, identified by positional cloning, that might be a disease-causing gene.

Capsid The protein coat that encloses the DNA or RNA molecule of a bacteriophage or virus.

Cassette A DNA sequence consisting of promoter-ribosome binding site-unique restriction site-terminator (or for a eukaryotic host, promoter-unique restriction site-polyadenylation sequence) carried by certain types of expression vector. A foreign gene inserted into the unique restriction site is placed under control of the expression signals.

Cauliflower mosaic virus (CaMV) The best studied of the caulimoviruses, used in the past as a cloning vector for some species of higher plant. Cauliflower mosaic virus is the source of strong promoters used in other types of plant cloning vector.

Caulimoviruses One of the two groups of DNA viruses to infect plants, the members of which have potential as cloning vectors for some species of higher plant.

Cell extract A preparation consisting of a large number of broken cells and their released contents.

Cell-free translation system A cell extract containing all the components required for protein synthesis (i. e. ribosomal subunits, tRNAs, amino acids, enzymes and cofactors) and able to translate added mRNA molecules.

Chimera ① A recombinant DNA molecule made up of DNA fragments from more than one organism, named after the mythological beast. ② The initial product of cloning using embryonic stem cells: an animal made up of a mixture of cells with different genotypes.

Chromosome One of the DNA-protein structures that contains part of the nuclear genome of a eukaryote. Less accurately, the DNA molecule(s) that contains a prokaryotic genome.

Chromosome walking A technique that can be used to construct a clone contig by identifying overlapping fragments of cloned DNA.

Cleared lysate A cell extract that has been centrifuged to remove cell debris, subcellular particles and possibly chromosomal DNA.

Clone A population of identical cells, generally those containing identical recombinant DNA molecules.

Clone contig approach A genome sequencing strategy in which the molecules to be sequenced are broken into manageable segments, each a few hundred kilobases or a few megabases in length, which are sequenced individually.

Clone fingerprinting Any one of a variety of techniques that compares cloned DNA fragments in order to identify ones that overlap.

Codon bias The fact that not all codons are used equally frequently in the genes of particular organism.

Combinatorial screening A technique that reduces the number of PCRs or other analyses that must be performed by combining samples in an ordered fashion, so that a sample giving a particular result can be identified even though that sample is not individually examined.

Compatibility The ability of two different types of plasmid to coexist in the same cell.

Competent A culture of bacteria that has been treated to enhance their ability to take up DNA molecules.

Complementary Two polynucleotides that can base pair to form a double-stranded molecule.

Complementary DNA (cDNA) cloning A cloning technique involving conversion of purified mRNA to DNA before insertion into a vector.

Conformation The spatial organization of a molecule. linear and circular are two possible conformations of a polynucleotide.

Conjugation Physical contact between two bacteria, usually associated with transfer of DNA from one cell to the other.

Consensus sequence A nucleotide sequence used to describe a large number of related though non-identical sequences. Each position of the consensus sequence represents the nucleotide most often found at that position in the real sequences.

Contig A contiguous segment of DNA sequence obtained as part of a genome sequencing project.

Continuous culture The culture of microorganisms in liquid medium under controlled conditions, with additions to and removals from the medium over a lengthy period of time.

Contour clamped homogeneous electric fields (CHEF) An electrophoresis technique for the separation of large DNA molecules.

Copy number The number of molecules of a plasmid contained in a single cell.

Cosmid A cloning vector consisting of the λ *cos* site inserted into a plasmid, used to clone DNA fragments up to 40kb in size.

***cos* site** One of the cohesive, single-stranded extensions present at the ends of the DNA molecules of certain strains of λ phage.

Covalently closed-circular DNA (cccDNA) A completely double-stranded circular DNA molecule, with no nicks or discontinuities, usually with a supercoiled conformation.

CpG island A GC-rich DNA region located upstream of approximately 56% of the genes in the human genome.

Denfined medium A bacterial growth medium in which all the components are known.

Deletion analysis The identification of control sequences for a gene by determining the effects on gene expression of specific deletions in the upstream region.

Denaturation Of nucleic acid molecules: breakdown by chemical or physical means of the hydrogen bonds involved in base pairing.

Density gradient centrifugation Separation of molecules and particles on the basis of buoyant density, by centrifugation in a concentrated sucrose or caesium chloride solution.

Deoxyribonuclease An enzyme that degrades DNA.

Dideoxynucleotide A modified nucleotide that lacks the 3′ hydroxyl group and so prevents further chain elongation when incorporated into a growing polynucleotide.

Direct gene transfer A cloning process that involves transfer of a gene into a chromosome without the use of a cloning vector able to replicate in the host organism.

Disarmed plasmid A Ti plasmid that has had some or all of the T-DNA genes removed, so it is no longer able to promote cancerous growth of plant cells.

DNA chip A wafer of silicon carrying a high density array of oligonucleotides used in transcriptome and other studies.

DNA marker A DNA sequence that exists as two or more alleles and which can therefore be used in genetic mapping.

DNA polymerase An enzyme that synthesizes DNA on a DNA or RNA template.

DNA profiling A PCR technique that determines the alleles present at different STR loci within a genome in order to use DNA information to identify individuals.

DNA sequencing Determination of the order of nucleotides in a DNA molecule.

Double digestion Cleavage of a DNA molecule with two different restriction endonucleases, either concurrently or consecutively.

Electrophoresis Separation of molecules on the basis of their charge-to-mass ratio.

Electroporation A method for increasing DNA uptake by protoplasts through prior exposure to a high voltage, which results in the temporary formation of small pores in the cell membrane.

Embryonic stem (ES) cell A totipotent cell from the embryo of a mouse or other organism, used in construction of a transgenic animal such as a knockout mouse.

End filling Conversion of a sticky end to a blunt end by enzymatic synthesis of the complement to the single-stranded extension.

Endonuclease An enzyme that breaks phosphodiester bonds within a nucleic acid molecule.

Episome A plasmid capable of integration into the host cell's chromosome.

Ethanol precipitation Precipitation of nucleic acid molecules by ethanol plus salt, used primarily as a means of concentrating DNA.

Ethidium bromide A fluorescent chemical that intercalates between base pair in a double-

stranded DNA molecule, used in the detection of DNA.

Exonuclease An enzyme that sequentially removes nucleotide from the ends of a nucleic acid molecule.

Expressed sequence tag (EST) A partial or complete cDNA sequence.

Expression vector A cloning vector designed so that a foreign gene inserted into the vector is expressed in the host organism.

Fermenter A vessel used for the large scale culture of microorganisms.

Field inversion gel electrophoresis (FIGE) An electrophoresis technique for the separation of large DNA molecules.

Fluorescence *in situ* hybridization (FISH) A hybridization technique that uses fluorochromes of different colours to enable two or more genes to be located within a chromosome preparation in a single *in situ* experiment.

Footprinting The identification of a protein binding site on a DNA molecule by determining which phosphodiester bonds are protected from cleavage by DNase I.

Functional genomics Studies aimed at identifying all the genes in a genome and determining their functions.

Gel electrophoresis Electrophoresis performed in a gel matrix so that molecules of similar electric charge can be separated on the basis of size.

Gel retardation A technique that identifies a DNA fragment that has a bound protein by virtue of its decreased mobility during gel electrophoresis.

Geminivirus One of the two groups of DNA viruses that infect plants, the members of which have potential as cloning vectors for some species of higher plants.

Gene A segment of DNA that codes for an RNA and/or polypeptide molecule.

Gene addition A genetic engineering strategy that involves the introduction of a new gene or group of genes into an organism.

Gene cloning Insertion of a fragment of DNA, carrying a gene, into a cloning vector, and subsequent propagation of the recombinant DNA molecule in a host organism. Also used to describe those technique that achieve the same result without the use of a cloning vector(e. g. direct gene transfer).

Gene knockout A technique that results in inactivation of a gene, as a means of determining the function of that gene.

Gene mapping Determination of the relative positions of different genes on a DNA molecule.

Gene subtraction A genetic engineering strategy that involves the inactivation of one or more of an organism's genes.

Gene therapy A clinical procedure in which a gene or other DNA sequence is used to treat a disease.

Genetic engineering The use of experimental techniques to produce DNA molecules containing new genes or new combinations of genes.

Genetic fingerprinting A hybridization technique that determines the genomic distribution of a hypervariable dispersed repetitive sequence and results in a banding pattern that is specific for each individual.

Genetic map A genome map that has been obtained by analysing the results of genetic crosses.

Genetics The branch of biology devoted to the study of genes.

Genome The complete set of genes of an organism.

Genomic library A collection of clones sufficient in number to include all the genes of a particular organism.

Genomics The study of a genome, in particular the complete sequencing of a genome.

GM (genetically modified) crop A crop plant that has been engineered by gene addition or gene subtraction.

Harvesting The removal of microorganisms from a culture, usually by centrifugation.

Helper phage A phage that is introduced into a host cell in conjunction with a related cloning vector, in order to provide enzymes required for replication of the cloning vector.

Heterologous probing The use of a labelled nucleic acid molecule to identify related molecules by hybridization probing.

Homology Refers to two genes from different organisms that have evolved from the same ancestral gene. Two homologous genes are usually sufficiently similar in sequence for one to be used as a hybridization probe for the other.

Homology search A technique in which genes with sequences similar to that of an unknown gene are sought, in order to confirm a gene identification or to understand the function of the unknown gene.

Homopolymer tailing The attachment of a sequence of identical nucleotides (e. g. AAAAA) to the end of a nucleic acid molecule, usually referring to the synthesis of single-stranded homopolymer extensions on the ends of a double-stranded DNA molecule.

Horseradish peroxidase An enzyme that can be complexed to DNA and which is used in a non-radioactive procedure for DNA labelling.

Host-controlled restriction A mechanism by which some bacteria prevent phage attack through the synthesis of a restriction endonuclease that cleaves the non-bacterial DNA.

Hybrid-arrest translation (HART) A method used to identify the polypeptide coded by a cloned gene.

Hybridization probe A labelled nucleic acid molecule that can be used to identify complementary or homologous molecules through the formation of stable base-paired hybrids.

Hybrid-release translation (HRT) A method used to identify the polypeptide coded by a cloned gene.

Hypervariable dispersed repetitive sequence The type of human repetitive DNA sequence used in genetic fingerprinting.

Immunological screening The use of an antibody to detect a polypeptide synthesized by a

cloned gene.

Inclusion body A crystalline or paracrystalline deposit within a cell, often containing substantial quantities of insoluble protein.

Incompatibility group Comprises a number of different types of plasmid, often related to each other, that are unable to coexist in the same cell.

Indel A position where a DNA sequence has been inserted into or deleted from a genome, so-called because it is impossible from comparison of two sequences to determine which alternative has occurred-insertion into one genome or deletion from the other.

Induction ① Of a gene: the switching on of the expression of a gene or group of genes in response to a chemical or other stimulus. ② Of λ phage: the excision of the integrated from of λ and accompanying switch to the lytic mode of infection, in response to a chemical or other stimulus.

Insertional inactivation A cloning strategy whereby insertion of a new piece of DNA into a vector inactivates a gene carried by the vector.

Insertion vector A λ vector constructed by deleting a segment of nonessential DNA.

***In situ* hybridization** A technique for gene mapping involving hybridization of a labelled sample of a cloned gene to a large DNA molecule, usually a chromosome.

Interspersed repeat element PCR (IRE-PCR) A clone fingerprinting technique that uses PCR to detect the relative positions of repeated sequences in cloned DNA fragments.

***In vitro* mutagenesis** Any one of several techniques used to produce a specified mutation at a predetermined position in a DNA molecule.

***In vitro* packaging** Synthesis of infective λ particles from a preparation of λ capsid proteins and a catenane of DNA molecules separated by *cos* sites.

Kinship analysis An examination of DNA profiles or other information to determine if two individuals are related.

Klenow fragment (of DNA polymerase I) A DNA polymerase enzyme, obtained by chemical modification of *E. coli* DNA polymerase I, used primarily in chain termination DNA sequencing.

Knockout mouse A mouse that has been engineered so that it carries an inactivated gene.

Labelling The incorporation of a marker nucleotide into a nucleic acid molecule. The marker is often but not always a radioactive or fluorescent label.

Lac selection A means of identifying recombinant bacteria containing vectors that carry the *lacZ′* gene. The bacteria are plated on a medium that contains an analogue of lactose that gives a blue colour in the presence of β-galactosidase activity.

Lambda (λ) A bacteriophage that infects *E. coli*, derivatives of which are used as cloning vectors.

Ligase (DNA ligase) An enzyme that, in the cell , repairs single-stranded discontinuities in double -stranded DNA molecules. Purified DNA ligase is used in gene cloning to join DNA

molecules together.

Linkage analysis A technique for mapping the chromosomal position of a gene by comparing its inheritance pattern with that of genes and other loci whose map positions are already known.

Linker A synthetic, double-stranded oligonucleotide used to attach sticky ends to a blunt-ended molecule.

Lysogen A bacterium that harbours a prophage.

Lysogenic infection cycle The pattern of phage infection that involves integration of the phage DNA into the host chromosome.

Lysozyme An enzyme that weakens the cell walls of certain types of bacteria.

Lytic infection cycle The pattern of infection displayed by a phage that replicates and lyses the host cell immediately after the initial infection. Integration of the phage DNA molecule into the bacterial chromosome dose not occur.

M13 A bacteriophage that infects *E. coli*, derivatives of which are used as cloning vectors.

Mapping reagent A collection of DNA fragments spanning a chromosome or the entire genome and used in STS mapping.

Melting temperature (T_m) The temperature at which a double-stranded DNA or DNA-RNA molecule denatures.

Messenger RNA (mRNA) The transcript of a protein-coding gene.

Microarray A set of genes or cDNAs immobilized on a glass slide and used in transcriptome studies.

Microinjection A method of introducing new DNA into a cell by injecting it directly into the nucleus.

Microsatellite A polymorphism comprising tandem copies of, usually, two-, three-, four- or five-nucleotide repeat units. Also called a short tandem repeat (STR).

Minimal medium A defined medium that provides only the minimum number of different nutrients needed for growth of a particular bacterium.

Mitochondrial DNA The DNA molecules present in the mitochondria of eukaryotes.

Modification interference assary A technique that uses chemical modification to identify nucleotides involved in interactions with a DNA-binding protein.

Multicopy plasmid A plasmid with a high copy number.

Multigene family A number of identical or related genes present in the same organism, usually coding for a family of related polypeptides.

Multiplex PCR A PCR carried out with more than one pair of primers and hence targeting two or more sites in the DNA being studied.

2μm circle A plasmid found in the yeast *Saccharomyces cerevisiae* and used as the basis for a series of cloning vectors.

Nick A single-strand break, involving the absence of one or more nucleotides, in a double-

stranded DNA molecule.

Nick translation The repair of a nick with DNA polymerase I, usually to introduce labelled nucleotides into a DNA molecule.

Northern transfer A technique for transferring bands of RNA from an agarose gel to a nitrocellulose or nylon membrane.

Nucleic acid hybridization Formation of a double-stranded molecule by base pairing between complementary or homologous polynucleotides.

Oligonucleotide A short, synthetic, single-stranded DNA molecule, such as one used as a primer in DNA sequencing or PCR.

Oligonucleotide-directed mutagenesis An *in vitro* mutagenesis technique that involves the use of a synthetic oligonucleotide to introduce the predetermined nucleotide alteration into the gene to be mutated.

Open-circular DNA (ocDNA) The non-supercoiled conformation taken up by a circular double-stranded DNA molecule when one or both polynucleotides carry nicks.

Open reading frame (ORF) A series of codons that is or could be a gene.

Origin of replication The specific position on a DNA molecule where DNA replication begins.

Orphan An open reading frame thought to be a functional gene but to which no function has yet been assigned.

Orthogonal field alternation gel electrophoresis (OFAGE) A gel electrophoresis technique that employs a pulsed electric field to achieve separation of very large molecules of DNA.

P1 A bacteriophage that infects *E. coli*, derivatives of which are used as cloning vectors.

P1-derived artificial chromosome (PAC) A cloning vector based on the P1 bacteriophage, used for cloning relatively large fragments of DNA in *E. coli*.

Papillomaviruses A group of mammalian viruses, derivatives of which have been used as cloning vectors.

Partial digestion Treatment of a DNA molecule with a restriction endonuclease under such conditions that only a fraction of all the recognition sites are cleaved.

Pedigree analysis The use of a human family tree to analyse the inheritance of a genetic or DNA marker.

P element A transposon from *Drosophila melanogaster* used as the basis of a cloning vector for that organism.

Phage display A technique involving cloning in M13 that is used to identify proteins that interact with one another.

Phage display library A collection of M13 clones carrying different DNA fragments, used in phage display.

Phagemid A double-stranded plasmid vector that possesses an origin of replication from a filamentous phage and hence can be used to synthesize a single-stranded version of a cloned gene.

Pharming Genetic modification of a farm animal so that the animal synthesizes a recombinant

pharmaceutical protein, often in its milk.

Physical map A genome map that has been obtained by direct examination of DNA molecules.

Pilus One of the structures present on the surface of a bacterium containing a conjugative plasmid, through which DNA is assumed to pass during conjugative.

Plaque A zone of clearing on a lawn of bacteria caused by lysis of the cells by infecting phage particles.

Plasmid A usually circular piece of DNA, primarily independent of the host chromosome, often found in bacteria and some other types of cells.

Plasmid amplification A method involving incubation with an inhibitor of protein synthesis aimed at increasing the copy number of certain types of plasmid in a bacterial culture.

Polyethylene glycol A polymeric compound used to precipitate macromolecules and molecular aggregates.

Polylinker A synthetic double-stranded oligonucleotide carrying a number of restriction sites.

Polymerase chain reaction (PCR) A technique that enables multiple copies of a DNA molecule to be generated by enzymatic amplification of a target DNA sequence.

Polymorphism Refers to a locus that is present as a number of different alleles or other variations in the population as a whole.

Positional cloning A procedure that uses information on the map position of a gene to obtain a clone of that gene.

Positional effect Refers to the variations in expression levels observed for genes inserted at different positions in a genome.

Post-genomics Studies aimed at identifying all the genes in a genome and determining their functions.

Primer A short single-stranded oligonucleotide which, when attached by base pairing to a single-stranded template molecule, acts as the start point for complementary strand synthesis directed by a DNA polymerase enzyme.

Primer extension A method of transcript analysis in which the 5′ end of an RNA is mapped by annealing and extending an oligonucleotide primer.

Promoter The nucleotide sequence, upstream of a gene, that acts as a signal for RNA polymerase binding.

Prophage The integrated form of the DNA molecule of a lysogenic phage.

Protease An enzyme that degrades protein.

Protein A A protein from the bacterium *Staphylococcus aureus* that binds specifically to immunoglobulin G (i. e. antibody) molecules.

Protein engineering A collection of techniques, including but not exclusively gene mutagenesis, that result in directed alterations being made to protein molecules, often to improve the properties of enzymes used in industrial processes.

Proteome The entire protein content of a cell or tissue.

Protoplast　A cell from which the cell wall has been completely removed.

RACE (rapid amplification of cDNA ends)　A PCR-based technique for mapping the end of an RNA molecule.

Radiation hybrid　A collection of rodent cell lines that contain different fragments of the human genome, constructed by a technique involving irradiation and used as a mapping reagent in studies of the human genome.

Radioactive marker　A radioactive atom used in the detection of a larger molecule into which it has been incorporated.

Random priming　A method for DNA labelling that utilizes random DNA hexamers, which anneal to single-stranded DNA and act as primers for complementary strand synthesis by a suitable enzyme.

Reading frame　One of the six overlapping sequences of triplet codons, three on each polynucleotide, contained in a segment of a DNA double helix.

Recombinant　A transformed cell that contains a recombinant DNA molecule.

Recombinant DNA molecule　A DNA molecule created in the test tube by ligating together pieces of DNA that are not normally contiguous.

Recombinant DNA technology　All of the techniques involved in the construction, study and use of recombinant DNA molecules.

Recombinant protein　A polypeptide that is synthesized in a recombinant cell as the result of expression of a cloned gene.

Recombination　The exchange of DNA sequences between different molecules, occurring either naturally or as a result of DNA manipulation.

Relaxed　Refers to the non-supercoiled conformation of open-circular DNA.

Repetitive DNA PCR　A clone fingerprinting technique that uses PCR to detect the relative positions of repeated sequences in cloned DNA fragments.

Replacement vector　A λ vector designed so that insertion of new DNA is by replacement of part of the non-essential region of the λ DNA molecule.

Replica plating　A technique whereby the colonies on an agar plate are transferred *en masse* to a new plate, on which the colonies grow in the same relative positions as before.

Replicative form (RF) of M13　The double-stranded form of the M13 DNA molecule found within infected *E. coli* cells.

Reporter gene　A gene whose phenotype can be assayed in a transformed organism, and which is used in, for example, deletion analyses of regulatory regions.

Repression　The switching off of expression of a gene or a group of genes in response to a chemical or other stimulus.

Restriction analysis　Determination of the number and sizes of the DNA fragments produced when a particular DNA molecule is cut with a particular restriction endonuclease.

Restriction endonuclease　An endonuclease that cuts DNA molecules only at a limited number

of specific nucleotide sequences.

Restriction fragment length polymorphism (RFLP) A mutation that results in alteration of a restriction site and hence a change in the pattern of fragments obtained when a DNA molecule is cut with a restriction endonuclease.

Restriction map A map showing the positions of different restriction sites in a DNA molecule.

Retrovirus A virus with an RNA genome, able to insert into a host chromosome, derivatives of which have been used to clone genes in mammalian cells.

Reverse transcriptase An RNA-dependent DNA polymerase, able to synthesize a complementary DNA molecule on a template of single-stranded RNA.

RFLP linkage analysis A technique that uses a closely linked RFLP as a marker for the presence of a particular allele in a DNA sample, often as a means of screening individuals for a defective gene responsible for a genetic disease.

Ribonuclease An enzyme that degrades RNA.

Ribosome binding site The short nucleotide sequence upstream of a gene, which after transcription forms the site on the mRNA molecule to which the ribosome binds.

Ri plasmid An *Agrobacterium rhizogenes* plasmid, similar to the Ti plasmid, used to clone genes in higher plants.

Reverse transcription-PCR (RT-PCR) A PCR technique in which the starting material is RNA. The first step in the procedure is conversion of the RNA to cDNA with reverse transcriptase.

Selectable marker A gene carried by a vector and conferring a recognizable characteristic on a cell containing the vector or a recombinant DNA molecule derived from the vector.

Selection A means of obtaining a clone containing a desired recombinant DNA molecule.

Sequence tagged site (STS) A DNA sequence whose position has been mapped in a genome.

Short tandem repeat (STR) A polymorphism comprising tandem copies of, usually, two-, three-, four- or five-nucleotide repeat units. Also called a microsatellite.

Shotgun approach A genome sequencing strategy in which the molecules to be sequenced are randomly broken into fragments which are then individually sequenced.

Shotgun cloning A cloning strategy that involves the insertion of random fragments of a large DNA molecule into a vector, resulting in a large number of different recombinant DNA molecules.

Shuttle vector A vector that can replicate in the cells of more than one organism (e. g. in *E. coli* and in yeast).

Simian virus 40 (SV40) A mammalian virus that has been used as the basis for a cloning vector.

Single nucleotide polymorphism (SNP) A point mutation that is carried by some individuals of a population.

Sonication A procedure that uses ultrasound to cause random breaks in DNA molecules.

Southern transfer A technique for transferring bands of DNA from an agarose gel to a nitro-

cellulose or nylon membrane.

Sphaeroplast A cell with a partially degraded cell wall.

Stem loop A hairpin structure, consisting of a base-paired stem and a non-base-paired loop, that may form in a polynucleotide.

Sticky end An end of a double-stranded DNA molecule where there is a single-stranded extension.

Strong promoter An efficient promoter that can direct synthesis of RNA transcripts at a relatively fast rate.

Stuffer fragment The part of a λ replacement vector that is removed during insertion of new DNA.

Supercoiled The conformation of a covalently closed-circular DNA molecule, which is coiled by torsional strain into the shape taken by a wound-up elastic band.

***Taq* DNA polymerase** The thermostable DNA polymerase that is used in PCR.

T-DNA The portion of the Ti plasmid transferred to the plant DNA.

Temperature-sensitive mutation A mutation that results in a gene product that is functional within a certain temperature range (e. g. at less than 30℃), but non-functional at different temperatures (e. g. above 30℃).

Template A single-stranded polynucleotide (or region of a polynucleotide) that directs synthesis of a complementary polynucleotide.

Terminator The short nucleotide sequence, downstream of a gene, that acts as a signal for termination of transcription.

5′terminus One of the two ends of a polynucleotide: that which carries the phosphate group attached to the 5′position of the sugar.

3′terminus One of the two ends of a polynucleotide: that which carries the hydroxyl group attached to the 3′position of the sugar.

Thermal cycle sequencing A DNA sequencing method that uses PCR to generate chain-terminated polynucleotides.

Ti plasmid The large plasmid found in those *Agrobacterium tumefaciens* cells able to direct crown gall formation on certain species of plants.

Total cell DNA Consists of all the DNA present in a single cell or group of cells.

Transcript analysis A type of experiment aimed at determining which portions of a DNA molecule are transcribed into RNA.

Transcriptome The entire mRNA content of a cell or tissue.

Transfection The introduction of purified virus DNA molecules into any living cell.

Transformation The introduction of any DNA molecule into any living cell.

Transformation frequency A measure of the proportion of cells in a population that are transformed in a single experiment.

Transgenic animal An animal that possesses a cloned gene in all of its cells.

Transposon A DNA sequence that is able to move from place to place within a genome.

Undefined medium A growth medium in which not all the components have been identified.

UV absorbance spectrophotometry A method for measuring the concentration of a compound by determining the amount of ultraviolet radiation absorbed by a sample.

Vector A DNA molecule, capable of replication in a host organism, into which a gene is inserted to construct a recombinant DNA molecule.

Vehicle Often used as a substitute for the word 'vector', emphasizing that the vector transports the inserted gene through the cloning experiment.

Virus chromosome The DNA or RNA molecule(s) contained within a virus capsid and carrying the viral genes.

Watson-Crick rules The base pairing rules that underlie gene structure and expression. A pairs with T, and G pairs with C.

Weak promoter An inefficient promoter that directs synthesis of RNA transcripts at a relatively low rate.

Western transfer A technique for transferring bands of protein from an electrophoresis gel to a membrane support.

Yeast artificial chromosome (YAC) A cloning vector comprising the structural components of a yeast chromosome and able to clone very large pieces of DNA.

Yeast episomal plasmid (YEp) A yeast vector carrying the 2μm circle origin of replication.

Yeast integrative plasmid (YIp) A yeast vector that relies on integration into the host chromosome for replication.

Yeast replicative plasmid (YRp) A yeast vector that carries a chromosomal origin of replication.

Yeast two hybrid system A technique involving cloning in *S. cerevisiae* that is used to identify proteins that interact with one another.

Zoo blot A nitrocellulose or nylon membrane carrying immobilized DNAs of several species, used to determine if a gene from one species has homologues in the other species.

(引自 Gene Cloning and DNA Analysis) (郭 春 曹英林)

附录Ⅷ 参考文献

1. Brown T.A.2001.Gene Cloning and DNA Analysis:An introduction.4th ed.Oxford:Blackwell Synergy Ltd
2. Dennis C, Gallagher R.2001.The Human Genome.1st ed.Hampshire:Palgrave Macmillan Ltd
3. Esmond S.T.Nicholl.2002.An introduction to Genetic Engineering.2ed ed.Cambridge:Cambridge University Press
4. Joseph Sambrook, David W.Russell.2000.Molecular Cloning:A Laboratory manual.3th ed.New York:Cold Spring Harbor Laboratory Press

5. Kaneta Y, Kagami Y, Tsunoda T, Ohno R, Nakamura Y, Katagiri T. 2003. Genome-wide analysis of gene-expression profiles in chronic myeloid leukaemia cells using a cDNA microarray. Int J Oncol, 23(3):681—691

6. R. Wold, S. B Primrose. 2001. Principles of Gene Manipulation. 6th ed. Oxford: Blackwell Synergy Ltd

7. Wang DN, Liu SY, Trummer BJ et al. 2002. Carbohydrate microarrays for the detection of recognization of cross-reactive molecular markers of microbes and host cells. Nature Biotechnology, 20:275—281

8. 贺林. 2001. 解码生命——人类基因组计划和后基因组计划. 北京:科学出版社

9. 卢圣栋. 1999. 现代分子生物学实验技术. 第2版. 北京:中国协和医科大学出版社

10. 吴乃虎. 2001. 基因工程原理. 第2版. 北京:科学出版社

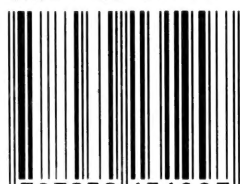




(R-1322.0101)







ISBN 7-03-013199-1

9 787030 131997 >

科学出版社 医学出版分社

E-mail: med_sp@sohu.com

ISBN 7-03-013199-1

定 价：32.00 元